

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra informatiky a výpočetní techniky

Diplomová práce

Statistické zpracování dat ruční antropometrie pro různé populace

ZÁPADOČESKÁ UNIVERZITA V PLZNI
Fakulta aplikovaných věd
Akademický rok: 2018/2019

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Jana KOUKALOVÁ**
Osobní číslo: **A16N0078P**
Studijní program: **N3902 Inženýrská informatika**
Studijní obor: **Informační systémy**
Název tématu: **Statistické zpracování dat ruční antropometrie pro různé populace**
Zadávací katedra: **Katedra informatiky a výpočetní techniky**

Z á s a d y p r o v y p r a c o v á n í :

1. Stručně popište problematiku ruční antropometrie pro potřeby návrhu ručního nástroje.
2. Na základě dat proveďte analýzu českých probandů a zjistěte, zda existují statisticky významné rozdíly mezi pohlavím, věkem a lateralitou. Dále zjistěte, zda existuje závislost mezi typem povolání a BMI (Body Mass Index). Algoritmy zpracování navrhňte tak, aby byla možná snadná aktualizace výpočtů pro nově naměřená data.
3. Porovnejte výsledky s dostupnými studiemi ruční antropometrie pro ostatní populace (například probandi Německa, Vietnamu, Turecka, atd.).
4. Navrhňte vhodnou vizualizaci výsledků pomocí webové aplikace.
5. Všechny výpočty doložte vhodnou programovou realizací. K práci přiložte datový nosič, který bude obsahovat všechny vytvořené programy a použitá data.

Rozsah grafických prací: **dle potřeby**
Rozsah kvalifikační práce: **doporuč. 50 s. původního textu**
Forma zpracování diplomové práce: **tištěná**
Seznam odborné literatury:
dodá vedoucí diplomové práce

Vedoucí diplomové práce: **RNDr. Blanka Šedivá, Ph.D.**
Katedra matematiky
Konzultant diplomové práce: **Ing. Marek Bureš, Ph.D.**
Regionální technologický institut

Datum zadání diplomové práce: **10. září 2018**
Termín odevzdání diplomové práce: **16. května 2019**

Radová

Doc. Dr. Ing. Vlasta Radová
děkanka



Brada

Doc. Ing. Přemysl Brada, MSc., Ph.D.
vedoucí katedry

Prohlášení

Prohlašuji, že jsem diplomovou práci vypracovala samostatně a výhradně s použitím citovaných pramenů a literatury.

V Plzni dne

Bc. Jana Koukalová

Poděkování

Ráda bych poděkovala vedoucí práce RNDr. Blance Šedivé, Ph.D., konzultantům z Regionálního technologického institutu Ing. Markovi Burešovi, Ph.D. a Ing. Věře Čadkové za vstřícný přístup, ochotu, trpělivost při zpracování celé práce. Dále bych ráda poděkovala všem, kteří mě během studia podporovali, rodině i přátelům.

Abstract

The purpose of this study was to create an useful tool providing statistical data processing using various methods. Its functionality was demonstrated by conducting a study dealing with hand anthropometry data of Czech population. Solution is created by web application using open-source technology R Shiny. It offers simple user interface outputting overview tables and graphs that are based on statistical methods. Information obtained from anthropometrical data reveals important facts about differences between groups. Results of this thesis serve as the source of information important for design of hand tools and requisites with emphasis for ergonomic shape.

Keywords: hand anthropometry, data analysis, statistical methods, statistical tool, web application, Shiny, R

Abstrakt

Účelem této práce bylo vytvořit nástroj pro statistické zpracování dat za pomoci různých metod. Jeho funkčnost byla demonstrována zpracováním studie zabývající se antropometrickými údaji rukou české populace. Programové řešení je realizováno pomocí webové aplikace využívající open-source technologii R Shiny. Ta poskytuje jednoduché uživatelské prostředí, jehož výstupem jsou přehledové tabulky a grafy, které jsou vytvořeny na základě statistických metod. Údaje získané z antropometrických dat odhalují důležité skutečnosti o rozdílech mezi skupinami. Výsledky této práce slouží jako zdroj informací důležitých pro návrh ručních, ergonomicky vhodných nástrojů a pomůcek.

Klíčová slova: antropometrie rukou, analýza dat, statistické metody, statistický nástroj, webová aplikace, Shiny, R

Obsah

1	Úvod	1
2	Antropometrie rukou	3
3	Metodologie	6
3.1	Problematika sběru, čištění a kvality dat	6
3.2	Předzpracování dat a odlehlá pozorování	7
3.2.1	Základní popisné charakteristiky	7
3.2.2	Grafické metody	9
3.2.3	Jednorozměrné metody	10
3.3	Statistické metody zpracování	11
3.3.1	Testy normality	11
3.3.2	Míra závislosti	12
3.3.3	Hodnocení průměru a rozptylu	14
3.3.4	Analýza rozptylu	17
3.3.5	Analýza hlavních komponent a shlukování	19
4	Programové řešení	21
4.1	Vypracování aplikace	21
4.2	Postup při vytvoření aplikace	22
4.2.1	Aktivace účtu na shinyapps.io	27
4.2.2	Nasazení aplikace	27
4.3	Omezení a možné problémy s vytvořenou aplikací	28
4.4	Testování	29
5	Analýza dat	31
5.1	Sběr dat	31
5.2	Předzpracování a čištění dat	33
5.2.1	Příprava datového souboru	33
5.2.2	Čištění dat	33
5.3	Analýza dat české populace	34
5.3.1	Základní charakteristiky antropometrických dat	36
5.3.2	Testy normality	38
5.3.3	Korelační analýza	39
5.3.4	Analýza rozptylu (ANOVA)	42
5.3.5	Metoda hlavních komponent a shlukování	45

5.4	Porovnání s ostatními populacemi	48
6	Závěr	51
	Literatura	53
A	Seznam použitých zkratk	61
B	Přílohy	63
B.1	Instalace R a RStudia	63
B.2	Spuštění aplikace na vlastním počítači	63
B.3	Uživatelská příručka	64
B.3.1	Příprava datového souboru	64
B.3.2	Načtení dat (sekce Data Loading)	65
B.3.3	Základní statistiky (sekce Basic Statistics)	65
B.3.4	Pokročilá statistika (sekce Advanced Statistics)	68
B.3.5	Porovnání s ostatními populacemi (sekce Comparison)	69
B.4	Přílohy na CD	72
B.5	Tištěné přílohy	73

1 Úvod

V dnešní době se společnost zajímá o analýzy ve výzkumu i v osobním životě. Denně je nasbíráno obrovské množství dat, řádově přes exabajty dat. Toto množství narůstá s používáním chytrých zařízení, senzorů přenášejících data o poloze, teplotě, srdečním tepu, vytvářením příspěvků na sociálních sítích, používáním komunikačních prostředků, sledování videí, ale i zadáváním údajů do různých aplikací a internetových obchodů. Při výrobě v továrnách je nashromážděno další množství dat. Dokonce i pouhým vyhledáváním v prohlížečích jsou vytvořena nová data. Jen Google se podílí na 77 % ze všech uskutečněných vyhledání, tj. 40 000 vyhledaných příspěvků za vteřinu. Lze si představit kolik dat je vyprodukováno každý den. Za rok 2017 toto množství bylo odhadnuto na 2,5 exabajtů denně [30]. V roce 2020 se předpokládá, že každý jedinec vytvoří 1,7 megabajtů za sekundu [16].

Data jsou převážně ukládána do datových skladů pomocí procesu ETL (angl. *extract, transform, load*). Nejprve je proveden výběr údajů, poté jsou transformovány tak, aby údaje byly validní a očištěné od neúplných, chybných či nekonzistentních hodnot. Nakonec jsou vloženy do datového skladu. Je nezbytné údajům rozumět a vědět, které hodnoty mohou být zadány "špatně". Příprava dat je velmi důležitá, neboť zahrnutí chybných hodnot zanesou do výpočtů velké chyby, které na první dojem nemusí být zřejmé. Proto je velmi důležité údaje očistit a sjednotit, aby bylo zamezeno neshodám a zkresleným výsledkům. Kvalita dat je podstatné téma, které by nemělo být přehlíženo. Když jsou data připravena, lze je analyzovat a extrahovat z nich informace pomocí technik dolování dat. Tento proces je znám jako *Data mining*. Zahrnuje analytické metody a algoritmy. Jsou používány statistické metody, neuronové sítě, klasifikační a regresní úlohy, shlukování, predikční modely, apod. Následná interpretace výsledků by měla zahrnovat i hodnocení zvolených metod. Metody by měly být testovány, kontrolovány, zda byly správně použity a vyhodnocovány, zda daný model je použit správně a vhodně. Případně by měl být použit jiný. Kontrola předpokladů modelů je mnohdy opomíjena, ovšem tento krok je klíčový ke správné interpretaci výsledků. Získané informace jsou transformovány na znalosti. Proto, pokud je prováděna libovolná analýza, měly by být tyto aspekty zohledněny. Existuje mnoho softwaru, webových i BI (angl. *Business Intelligence*) řešení. Je ovšem nutné zvážit, zda možnosti dané aplikace jsou relevantní a správně použity.

Práce je zaměřena na problematiku analýzy dat antropometrie rukou, tedy oborem zabývajícím se měřením a popisem rozměrů ruky a dlaně. Tyto informace jsou důležité zejména při návrhu bezpečnostních i ergonomických nástrojů pro práci a přispívají k vývoji návrhu pokynů pro nástroje a pomůcky. Může se jednat například o rukojeť zbraně, ergonomickou myš či pracovní rukavice. Při návrhu ručních nástrojů je důležité zvážit antropometrii ruky, neboť nesprávná velikost může způsobit přetěžování šlach. Například ruční nástroj navržený pro muže je pravděpodobně větší než pro ženy. Ženy tedy budou natahovat prsty a namáhat jiné části ruky než by měly, a tak jejich šlachy budou častěji napnuté [33]. To může zapříčinit různá onemocnění, jako například syndrom karpálního tunelu či jiné onemocnění pohybového aparátu. Práce se bude zabývat danými rozměry ruky a bude zkoumáno, zda existují rozdíly rozměrů ruky pro různé kategorie. Následně budou porovnávány s jinými populacemi. Pro tyto účely bude realizován vhodný statistický nástroj.

Práce bude členěna do několika kapitol. Nejprve bude v práci stručně popsána problematika antropometrie rukou. Dále budou popsány statistické metody, které budou naimplementovány do statistického nástroje. Tato kapitola rovněž zmiňuje problematiku sběru, čištění a kvality dat. Další část práce tvoří popis použitého programového řešení, kde bude mimo jiné ověřena i kontrola správnosti vytvořených algoritmů. Poslední kapitola se zabývá analýzou dat antropometrie rukou pro českou populaci. Důležitou částí v této kapitole je i porovnání s ostatními národy.

2 Antropometrie rukou

V první kapitole je charakterizován pojem antropometrie rukou. Je vysvětleno, proč je důležité se zabývat problematikou antropometrie rukou pro návrh ručních nástrojů a pomůcek a kde se využívá.

Antropometrie, z řeckého *antropos*, je věda zabývající se živou bytostí (člověkem), konkrétně měřením, popisem a rozbořením tělesných znaků. Mezi tyto znaky patří například výška, váha, síla a velikost těla. Antropometrická data jsou důležitá, zejména proto, že poskytují informace pro různé studie, pro vývoj zařízení, vybavení a nástrojů.

Práce je věnována problematice antropometrie rukou, tj. tělesným rozměrům rukou. Lidská ruka je komplexní systém skládající se z prstů, článků prstů, záprstních i zápěstních kůstek, kloubů, šlach, vazů, nervů, cév a svalů, které mohou být poničeny či namoženy v důsledku nesprávného používání ručního nástroje či opakovaného namáhání. Jedná se o dokonalý nástroj, který je používán denně k nejrůznějším úkonům (např. držení předmětu, podání předmětu, otáčení zápěstí, mačkání tlačítek a kláves, hnětení). Bez rukou by nebylo vyrobeno dalších nástrojů usnadňující práci, ať už se jedná o práci fyzicky náročnou, manuální či kancelářskou. Sportovci hrající raketové či míčové sporty, chirurgové, hudebníci, kadeřníci, maséři, piloti, zedníci, programátoři, vědci, umělci, zemědělci, zkrátka všichni používají ke své práci ruce. Měly by být chráněny a používány co možná nejkvalitnější ergonomické nástroje, aby ruce sloužily co možná nejdéle. I odbornou veřejností neuznávaná chiromantie (čtení z ruky) rozlišuje čtyři typy rukou v závislosti na tvaru a velikosti ruky. Určují se z výpočtu délky ruky, velikosti prostředníku, šířky a délky dlaně [43].

Antropometrie rukou může být nápomocná i při forenzní analýze - identifikace pohlaví osob ze zbylých částí těla [27]. Nemusí se ovšem vždy jednat o návrh nástroje, někdy stačí danému nástroji snížit hmotnost. Ve výzkumu týkajícího se vlivu hmotnosti sbíječky na svalovou aktivitu, porovnávali konvenční 41 kg sbíječku s lehčí 27 kg, přičemž u lehčí varianty zaznamenali snížení námahy svalové aktivity [26]. Aby ergonomové a výzkumníci navrhli co nejlepší a nejpřesnější model, mohou využít trojrozměrná antropometrická data, která získají při 3D skenování rukou [42].

Nejrůznější studie poukazují na zdravotní problémy, potíže a poruchy pro manuálně pracující populaci. Jedná se zejména o poranění pohybového aparátu [33], poranění šlach a poškození z opakovaného namáhání (angl. *Repetitive Strain Injury*, zkr. RSI, či *Cumulative Trauma Disorder*, zkr. CTD)

[6]. Nejčastější příčina RSI je přetížení zápěstí, šlach a vazů při kancelářské i manuální práci s ergonomicky nevhodnými nástroji. Syndromem karpálního tunelu trpí 3,8 % celkové populace, tzn. každý pětadvacátý, a postihuje ženy třikrát častěji než muže [41]. Vzniká zvýšeným tlakem v karpálním tunelu a tím poškodí střední nerv (např. při úrazu zápěstí nebo během jeho přetěžování). Ovšem i jiné poškození šlach, vazů, zápěstí či hybnosti lokte a ramene negativně ovlivní náš pracovní výkon (tenisový loket). Jedna třetina nehod v průmyslovém odvětví postihuje funkcionalitu rukou, z toho jedna polovina se potýká s trvalými následky, která může vyústit v nefunkčnost ruky [9].

Proto se ergonomové, designéři, výzkumníci a konstruktéři ručních nástrojů či pracovních pomůcek snaží vyvíjet, vyrábět a kontrolovat své výrobky, aby splňovaly požadavky na bezpečnost práce a tělesnou konstrukci zákazníků. Může se jednat o návrh rukojeti, úchytu, počítačové myši, nůžek (zahradnických, elektrických, kancelářských, kadeřnických), ochranných pomůcek (pracovní rukavice, antivibrační rukavice, rukavice na žehlení vlasů, chirurgické rukavice, tepelně odolné rukavice, chrániče prstů, rukavice pro záchranáře a hasiče), kuchyňských a domácích pomůcek (mixér, nože, náprstek), sportovních pomůcek (rukavice, bandáže na zápěstí, ortézy, tejpovací pásky), řemeslných nástrojů (kladiva, kleště, šroubováky) či pracovních stanic v lehkém průmyslu. Další využití je ve vojenském vybavení letounů, vozidel, ponorek a tanků [14]. S vhodnými ergonomickými nástroji roste i efektivita a produktivita práce.

K ručním nástrojům a pomůckám je důležité znát kromě rozměrů ruky, i pohlaví, fyzickou zdatnost, zručnost, predispozice, národnost, zdravotní stav, věkovou skupinu či lateralitu rukou (dominance ruky - pravák, levák, ambidextrie). Tyto faktory mají vliv na návrh ručních nástrojů a pomůcek.

Při návrhu rukojeti je rozlišována válcová rukojeť, která snižuje výskyt RSI, eliptická snižující námahu a kontaktní plochu prstu a jiné rukojeti, které se zabývají otázkou velikosti a tvaru ručního nástroje. Rovněž se rozlišují i zásady pro návrh nástrojů. Průměr rukojeti okolo 30 až 40 mm je vhodný pro návrh zbraně či elektrického nástroje, ovšem mělo by platit, aby průměr nepřekročil jednu pětinu délky ruky. Průměrná délka rukojeti by se měla pohybovat mezi 100 až 125 mm, sklon okolo 78°. [45]

U manuálně pracující populace je nejvíce poruch pohybového aparátu. Prvním příkladem jsou krejčí a ševci. V rozvojových zemích se snaží oděvní fabriky minimalizovat náklady na výrobu oděvu a obuvi, proto využívají levnou pracovní sílu [39]. Ačkoliv denně ručně šijí, stříhají a sešívají látky, nemají adekvátní nástroje [6]. Dalším příkladem může být operátor výrobní linky. Pracovník na výrobní lince sestavuje výrobek z menších dílů. Proto je

důležité, aby jeho pracovní stanice splňovala ergonomické požadavky. Další skupinou jsou zahradníci, zemědělci, floristi, dřevorubci a další, kteří rekultivují krajinu. Většina kolumbijských květinářů ve studii [20] potvrdila, že používají nástroje, které neodpovídají jejich rozměrům ruky, což může způsobit neobvyklou mechanickou zátěž. Designéři se mohou takto zaměřit i na návrh zahradnických a zemědělských pomůcek (nůžky, pily, zastřihovače, křovinořezy, apod.).

Ovšem nemusí se jednat pouze o povolání. Kdokoliv, kdo píše textové zprávy, e-maily na mobilních telefonech, hraje hry a používá u toho joystick či gamepad, namáhá klouby a šlachy u palců. Syndrom z přetížení či bolavý palec může vyústit v syndrom karpálního tunelu. Proto i samotný návrh mobilních telefonů by měl brát ohledy na správnou velikost displejů. Někteří uživatelé s dlouhými prsty mají problém napsat SMS a stěžují si na malou velikost mobilních zařízení. Naopak lidé s kratšími palci a užšími dlaněmi si stěžují na držení většího telefonu. Proto rozvřzení klávesnice je pro uživatele mobilních telefonů klíčové. [8]

Pro hráče počítačových her, je důležitá ergonomická klávesnice, počítačová myš, i samotný joystick (pákový ovladač). Piloti rovněž ovládají stroje pomocí řídicí páky (kniplu). Především piloti stíhaček a jiných nejnovějších letounů s digitálním řízením ovládají letoun joystickem (namáhají palec, kterým ovládají kloboučkové tlačítko). Joysticky se ovšem objevují i u průmyslových strojů (jeřáby, vysokozdvizné plošiny). Ovšem prsty a šlachy se unaví i během trénování a hraní na hudební nástroje. Proto mohou někteří umělci trpět bolestmi.

Další kategorií jsou sportovci. Existuje mnoho pomůcek i ochranných prvků chránících před zlomeninami, naraženinami či opotřebením šlach, kloubů a vazů. Může se jednat o chrániče na zápěstí, boxovací rukavice, plavecké packy pro zesílení rukou, bandáže na zpevnění zápěstí, rukavice pro vzpírání, ergonomická držadla, berany ale i o trenažéry na posílení rukou (handbike trenažér). Jeden špatný úchop či špatná rotace prstu může způsobit zdravotní problém. Některé studie rovněž poukazují na závislost mezi délkou prstů a silou stisku [18]. Může ovšem i záležet na rozdílu dominantní a nedominantní ruky, protože nedominantní ruka bývá slabší.

3 Metodologie

Tato kapitola shrnuje metody zpracování dat a statistické testy. V práci jsou použita obecně známá názvosloví, jako například proband (jednotlivec, který je zkoumán), dataset (datový soubor, sada dat), populace (určitý soubor jedinců, který je předmětem zkoumání), populační skupina (skupina se společnými rysy, např. věková skupina) či laterita rukou (dominance rukou).

3.1 Problematika sběru, čištění a kvality dat

Při práci s daty je třeba si uvědomit, že převáděním dat, duplikací hodnot, načtením dat a sběru dat může nastat několik problémů. Aplikace nečekaně ukončí svůj běh, spočítají se nesmyslné hodnoty, zkreslí se výsledky, apod. Ačkoliv tato problematika nebývá hlavním cílem analýzy, krátce je v této sekci popsáno, co tyto události může způsobit.

U sběru dat záleží na tom, jakým způsobem byla získána. Pokud data byla získána ručně, existuje větší pravděpodobnost, že do dat je zanesena informační či výpočetní chyba. Může se jednat o překlep dotyčnou osobou, nečitelný rukopis nebo chybu v desetinné čárce. Pokud jsou data získána strojově (např. exportem z provozních systémů, replikací dat, přímým napojením na software), je třeba pamatovat na konzistenci dat (převedení datových typů, formátů, porušení pravidel omezení, nestejně pojmenování), duplikaci dat (při znovunačtení, angl. *recovery*), chyb při načítání dat do souboru (výpadek elektrického proudu), správnosti a úplnosti údajů (při archivaci a zálohování dat), kalibrací a přesností měřidla (v případě, jsou-li data sesbírána ze senzorů či pomoci zařízení). Je třeba tato data zkontrolovat a ověřit jejich správnost.

Před analýzou dat je tedy důležité data očistit a to buď ručně, nebo pomocí algoritmů a metod, jež jsou k dispozici. Může se jednat o hodnoty, které uživatelem nebyly vyplněny (neúplné hodnoty). Takovéto hodnoty jsou buďto zanedbány, odhadnuty (např. průměrem podobných skupin, predikcí) či označeny jako neznámé hodnoty (zpravidla jako 'NA'). Dále se může jednat o hodnoty, jež obsahují chybné hodnoty, které lze zjistit během statistické analýzy. Nicméně většinu z těchto chyb je třeba opravovat ručně.

Kvalitu dat lze posoudit pomocí metrik. V zásadě jsou k dispozici dvě metriky, objektivní a subjektivní. Jak z názvu vyplývá, *objektivní metriky* se zabývají statistickými a výpočetními charakteristikami (jsou číselné). Zatímco *subjektivní metriky* hodnotí způsob, jakým data vznikla (kvalita zdroje

dat, důvěryhodnost dat). Jedná se o subjektivní hodnocení na základě předchozích znalostí a zkušeností nebo vyjádření experta. Panuje snaha o to, aby se postupy standardizovaly pomocí norem, předpisů a zákonů, a tak nedocházelo k chybám. Data by měla být relevantní (míra toho, jak moc data splňují účel), přesná, aktuální (není třeba aktualizace), dostupná (problematika autorských práv, ochrana osobních údajů viz GDPR), porovnatelná s ostatními zdroji, koherentní (otázka postupu při sběru a vytvoření dat), úplná (jsou všechna data nahrána, jsou data rovnoměrně rozdělena), jedinečná (neobsahují duplicitní hodnoty) a konzistentní (je dodržen datový formát a typ, data souhlasí s fakty).

3.2 Předzpracování dat a odlehlá pozorování

Jsou-li data načtená, je vhodné zjistit, zda neobsahují nějaké netypické hodnoty. Odlehlá pozorování (angl. *outliers*) jsou nejčastějším případem. Tyto pozorování jsou chápány jako hodnoty, které se významně liší od jiných, vybočují ze zkoumaného vzorku, a tím ovlivňují následující výpočty a analýzu dat. Při sběru dat mohou být zaneseny chyby zapříčiněné lidským faktorem, náhodné chyby či chyby nastavení měřidel. Nepatrná změna v desetinné čárce má například velký dopad na statistické vyhodnocení zpracovávaných dat. Odlehlá pozorování mohou být i relevantní, pak ukazují na specifickou část zkoumaného vzorku. Proto existuje několik metod, jak tyto odlehlá pozorování najít.

3.2.1 Základní popisné charakteristiky

Prvním krokem při zkoumání dat jsou popisné charakteristiky. Minimum, maximum, aritmetický průměr, medián, modus, rozptyl a směrodatná odchylka. Dále standardně používané statistiky jsou koeficienty špičatosti a šikmosti, percentily, variační koeficient, směrodatná chyba průměru či směrodatná chyba 5. a 95. percentilů. Základní charakteristiky upozorňují na určité neshody, které mohou být v datech spatřeny. Například jak moc se liší minimální a maximální hodnota od průměru, kolikrát se průměrná hodnota nejčastěji opakovala (modus), zda rozptyl není moc velký, apod. Vzhledem k tomu, že při tvorbě antropometrické databáze je třeba dodržet jistý postup a měly by být uváděny dané statistické údaje [19]. Vyjma běžných základních charakteristik je třeba uvádět i další charakteristiky, které jsou uvedeny níže. Popis ostatních charakteristik je k dispozici v uživatelské příručce viz B.3.3.

Směrodatná chyba průměru, značena $s_{\bar{x}}(x)$, je definována následovně

$$s_{\bar{x}}(x) = \sqrt{\frac{s^2(x)}{n}}, \text{ kde} \quad (3.1)$$

x ... pozorování, $x \in R$

n ... počet pozorování

s^2 ... výběrový rozptyl.

Směrodatná chyba směrodatné odchylky, značena $s_{s(x)}(x)$, je definována jako

$$s_{s(x)}(x) = \sqrt{\frac{s^2(x)}{2n}}. \quad (3.2)$$

Směrodatná chyba 5. a 95. percentilů, značená $s_p(x)$, je definována následovně

$$\begin{aligned} s_p(x) &= \sqrt{s_{\bar{x}}^2(x) + 1,645^2 \cdot s_{s(x)}^2(x)} = \sqrt{\frac{s^2(x)}{n} + 1,645^2 \cdot \frac{s^2(x)}{2n}} = \\ &= \sqrt{\frac{s^2(x)}{n} \left(1 + \frac{1,645^2}{2}\right)} = \sqrt{\frac{s^2(x)}{n}} \cdot 2,353 = \\ &= \frac{s(x)}{\sqrt{n}} 2,353. \end{aligned} \quad (3.3)$$

V neposlední řadě **variační koeficient**, značen VK , je definován jako podíl výběrové směrodatné odchylky $s(x)$ a průměru \bar{x} vynásoben 100.

$$VK = \frac{s(x)}{\bar{x}} \cdot 100 \quad (3.4)$$

Někdy je užitečné data rozložit na intervaly. Pomocí **kvantilů, percentilů a kvartilů** lze zjistit, zda nejsou hodnoty od sebe výrazně rozptýleny. Je uveden rozdíl mezi těmito pojmy. **Kvantil**, značen x_q , je hodnota, jež definuje, kolik údajů leží pod jeho hladinou q . Kvantil, jehož hladina je uváděna v procentech, je označován jako **percentil** (např. kvantil o hladině 0.05 je 5% percentil). Existuje 99 možných percentilů (100 intervalů, každý obsahuje 1 % hodnot z celkové množiny hodnot) [21]. Percentily rozdělené po 25% hladině jsou nazývány **kvartily**. Dělí tedy 100% interval na 4 části. První (dolní) kvartil odpovídá 25% percentilu, druhý kvartil (medián) je 50% percentil a třetí (horní) 75% percentil. Percentily s hladinami 5 % a 95 % popisují krajní hodnoty rozdělení. [23]

Koeficienty špičatosti a šikmosti jsou další popisné charakteristiky. Pomocí nich lze zjistit, zda data odpovídají normálnímu rozdělení. V tomto

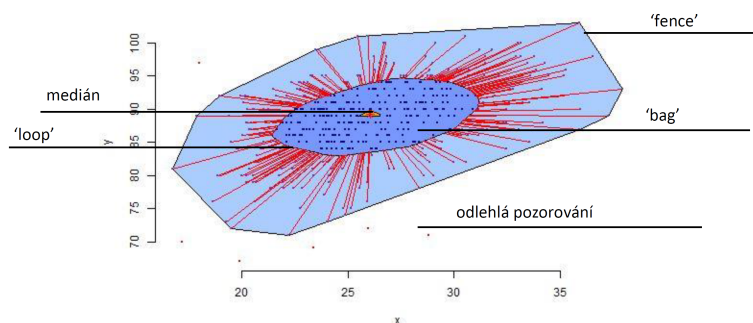
případě je koeficient šikmosti S nulový a koeficient špičatosti K je roven 3 (tvar křivky je symetrický). Data mohou být zešikmená doprava ($S > 0$), doleva ($S < 0$), tvar křivky je špičatější ($K < 3$, leptokurické) nebo méně špičatý ($K > 3$, platykurické) než u normálního rozdělení. Oba koeficienty vycházejí z centrálních momentů. Vzhledem k tomu, že vzorec momentu obsahuje průměr a u zešikmených dat se liší průměr a medián, zavedl Pearson svou míru šikmosti. Statistiky lze nalézt v literatuře viz [23].

3.2.2 Grafické metody

Dalším krokem při studiu dat jsou grafické metody. Pro svou názornost a intuitivnost je dospěno výsledku snadno. Mezi grafické metody detekující odlehlá pozorování patří vykreslení boxplotu (krabicového diagramu) nebo bagplotu (dvourozměrné verze boxplotu). Vykreslení boxplotu se používá, pokud je porovnáváno několik skupin. Bližší popis boxplotu je k nalezení v příručce viz B.3.3.

Bagplot je dvourozměrná podoba boxplotu, která se skládá, podobně jako boxplot, z několika částí (ilustrace viz obr. 3.1). První z nich je tzv. 'bag' (vak), v případě boxplotu se jedná o krabičku, která obsahuje 50 % hodnot. Střed tohoto 'bagu' tvoří tzv. Tukey medián (nejhlubší pozice v datech). Dále se skládá z tzv. fencu (plotu, ohraničení), jež je analogií horní a dolní antény v případě boxplotu (v bagplotu není zakreslen). Prostor, který vymezuje 'fence' a 'bag' se nazývá tzv. 'loop' (smyčka), jež reprezentuje vousy v boxplotu. Vzhledem k tomu, že se jedná o dvourozměrnou podobu boxplotu, je sledována variabilita dat (rozpětí bazu), korelace mezi proměnnou X a Y (natočení bazu), vychýlenost dat (tvar a vzdálenost bazu od smyčky, eliptický tvar značí vyváženost) a tzv. tails (ocasy, body blízko hranice mezi smyčkou a odlehlými pozorováními). Konstrukce bagplotu a bližší detaily o bagplotu jsou k nalezení v literatuře viz [38].

Obrázek 3.1: Popis bagplotu

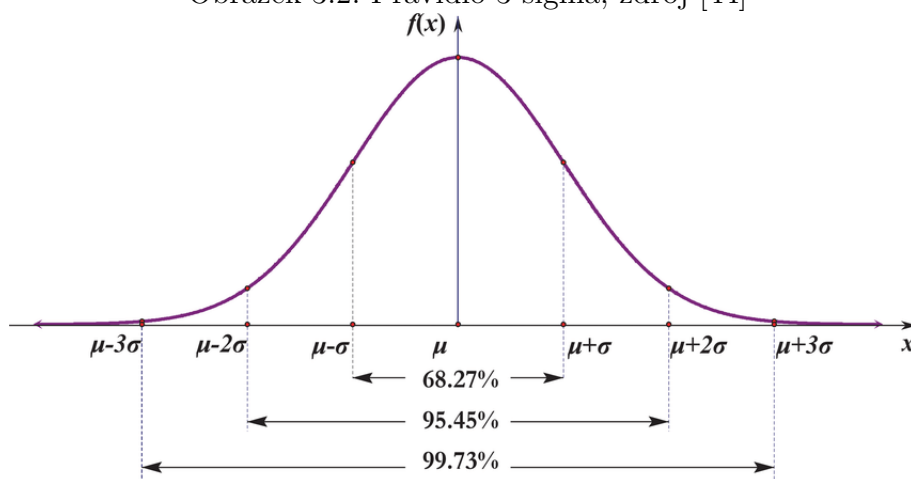


3.2.3 Jednorozměrné metody

Jednorozměrné metody jsou nejčastěji používané metody pro jejich jednoduchost. Předpokládají vztah rozdělení dat s nějakým předem známým rozdělením, obvykle normálním rozdělením. Nejznámější metodou v tomto případě je pravidlo 3-sigma.

Pravidlo 3-sigma říká, že pokud se předpokládá normální rozdělení dat, téměř všechny hodnoty leží ve vzdálenosti tří směrodatných odchylek od průměru. V případě jedné směrodatné odchylky od průměru se nachází 68,27 % hodnot dat ve vzdálenosti jedné směrodatné odchylky od průměru. Ve vzdálenosti dvou směrodatných odchylek leží 95,45 % hodnot a ve vzdálenosti tří směrodatných odchylek 99,73 % hodnot dat od průměru. Hodnoty, které leží mimo interval vzdáleností tří směrodatných odchylek od průměru, lze považovat za odlehlá pozorování.

Obrázek 3.2: Pravidlo 3-sigma, zdroj [44]



Normované hodnoty také mohou poukázat na vlivná odlehlá pozorování. Standardizací lze dosáhnout nulového průměru a jednotkové směrodatné odchylky. Hodnoty jsou normalizovány pomocí vzorce:

$$z_i = \frac{x_i - \bar{x}}{s(x)} \quad (3.5)$$

Data, která mají symetrické rozdělení, se po normalizaci pohybují v rozmezí od -3 do 3. Hodnoty mimo tento interval se považují za odlehlá pozorování. [22]

Ovšem lze použít i jiné metody pro identifikaci odlehlých pozorování, pro který předpoklad normality nemusí být splněn. Jednou z nich je **mediánová absolutní odchylka** (angl. *median absolute deviation*, zkr. MAD).

Tato statistika představuje míru variability a je rezistentní na odlehlá pozorování (robustní metoda), protože odečítá hodnoty od mediánu a ne od průměru. Obdobně jako u pravidla 3-sigma je vytvořen horní a dolní interval ve vzdálenosti tří směrodatných odchylek od mediánu ($\tilde{x} \pm 3 \cdot \hat{\sigma}$). Směrodatnou odchylku odhaduje pomocí MAD. Hodnoty mimo tento interval jsou klasifikovány jako odlehlá pozorování.

$$MAD = med(|x - \tilde{x}|) \quad (3.6)$$

$$\begin{aligned} \hat{\sigma} &= l \cdot MAD = \frac{-1}{erfc^{-1}(3/2) \cdot \sqrt{2}} MAD = \\ &= 1.4826 \cdot med(|x - \tilde{x}|), kde \end{aligned} \quad (3.7)$$

\tilde{x}, med	...	medián
$\hat{\sigma}$...	neparametrický odhad směrodatné odchylky
l	...	škálovací konstanta
$erfc$...	komplexní chybová funkce.

3.3 Statistické metody zpracování

V této sekci jsou shrnuty statistické metody, které byly použity ve vytvořené aplikaci. Zároveň jsou uvedeny předpoklady a vhodné použití těchto metod.

3.3.1 Testy normality

Předpoklad normality je velmi častým u mnoha statistických testů. Toto rozdělení, nazývané též Gaussovo, se v praxi často nevyskytuje, ale celkem dobře se pomocí tohoto rozdělení modelují proměnné. Je-li k dispozici mnoho nezávislých náhodných proměnných, jejich součet s rostoucím počtem pozorování odpovídá normálnímu rozdělení. Tuto domněnku vyjadřuje centrální limitní věta, která ukazuje, že mnoho rozdělení se asymptoticky blíží normálnímu rozdělení při větším počtu pozorování ($n > 30$, viz [23]). Má-li proměnná normální rozdělení, je zapisována jako $N(\mu, \sigma^2)$. Střední hodnota μ je odhadována průměrem, rozptyl výběrovým rozptylem $s^2(x)$.

Lze použít mnoho testů normality, každý test má své výhody a nevýhody. Ovšem nelze říci, který test je věrohodnější než jiný. V této části budou popsány některé z těchto testů a bude uvedeno, pro jaký počet pozorování se tyto testy využívají. Bližší popis testů lze najít v literatuře viz [37], [23], [7].

Jarqueho-Berův test vychází z koeficientu špičatosti a šikmosti. Pokud se jedná o normální rozdělení, koeficient špičatosti je roven 3 a koeficient šikmosti je roven 0. Využívá se pro větší počet pozorování $n \geq 200$. Ovšem obsahují-li data odlehlá pozorování, statistika není tolik robustní.

Sharipův-Wilkův test testuje, zda body v Q-Q grafu se liší od regresní přímky proložené těmito body. Je intuitivní, že v případě odlehlých hodnot je vzdálenost od proložené přímky a pozorování větší a tím může být zkreslen výsledek testu. Používá se pro menší počet pozorování, $n < 50$.

Kolmogorovův-Smirnovův test je vhodný pro jakékoliv rozdělení. Patří mezi testy dobré shody a testuje, zda distribuční funkce dat $F(x)$ pochází z rozdělení teoretické distribuční funkce $F_0(x)$. Testovací statistika je maximum z absolutního rozdílu empirické distribuční funkce a teoretické distribuční funkce. Určuje, jak moc daleko od sebe je empirická a teoretická distribuční funkce (v případě odlehlých pozorování vytvoří velký skok). Využívá se pro hodnoty $n \geq 20$. [23]

Lilieforsův test vychází z Kolmogorova-Smirnova testu, nicméně má upravenou kritickou mez, protože testuje data na normální rozdělení. Využívá se také při větším počtu pozorování $n \geq 100$.

Lze využít i grafické metody k zjištění, zda se data řídí normálním rozdělením. Nejčastěji používané je vykreslení histogramu, odhadnuté funkce hustoty na základě jádrového odhadu a Q-Q grafu. Bližší popis v uživatelské příručce viz B.3.3.

3.3.2 Míra závislosti

Pro zjištění vztahu mezi dvěma proměnnými lze použít korelaci či kovarianci. Pro zjištění toho, o jaký vztah mezi dvěma proměnnými se jedná (lineární, kvadratický, apod.), lze jednu proměnou odhadnout pomocí druhé. V tomto případě se jedná o regresní analýzu.

Nejprve je vhodné data vykreslit, protože tím je získána představa o tom, zda existuje mezi proměnnými určitá závislost či nikoliv. Dále je body proložena přímka, která je sestrojena pomocí metody nejmenších čtverců. Lze tak zodpovědět otázku, zda existuje lineární vztah mezi dvěma proměnnými. Jestliže jsou proměnné členěny podle několika kategorií, měly by se vyhodnocovat závislosti pro každou skupinu zvlášť. Nehomogenita dat může zapříčinit zkreslené výsledky.

Pro rozpoznání korelace bylo navrženo několik korelačních koeficientů. Jedná se zejména o Pearsonův, Spearmanův a Kendallův korelační koeficient. V některých případech mohou korelace dvou hodnot popisovat vztah, který

mezi těmito proměnnými neexistuje. Je třeba odchytil, zda korelace není způsobena klamným faktorem. Tato korelace je označována i jako nesmyslná či způsobená společnou příčinou. Příkladem může být vztah mezi délkou pravé a levé ruky. Klamné korelace jsou ovlivněny ještě třetí proměnou, v případě rukou se může jednat o výšku či o BMI (*body mass index*). V tomto případě je vhodné použít výběrový koeficient parciální korelace.

Pearsonův korelační koeficient

Tento koeficient vyjadřuje, zda existuje lineární závislost mezi dvěma proměnnými. Nabývá hodnot mezi -1 (nepřímá závislost) až 1 (přímá závislost). Hodnota korelačního koeficientu blízko nuly značí, že mezi proměnnými není lineární vztah. Je zkoumán vztah mezi kovariancí a směrodatnými odchylkami dvou proměnných. Korelační koeficient je tedy citlivý na odlehlá pozorování a vyjadřuje sílu pouze lineárního vztahu.

$$r_{xy} = \frac{s_{xy}}{s(x) \cdot s(y)} = \frac{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s(x) \cdot s(y)}, kde \quad (3.8)$$

r_{xy}	...	výběrový korelační koeficient
s_{xy}	...	výběrová kovariance
$s(x), s(y)$...	výběrové směrodatné odchylky proměnných.

Výběrová kovariance je definována vztahem

$$s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (3.9)$$

viz [37]. Kovarianci složek proměnných lze zapsat do kovarianční matice, kdy na diagonále jsou rozptyly jednotlivých proměnných.

Neparametrické varianty korelačních koeficientů

Další dva koeficienty, Spearmanův koeficient pořadí a Kendallův koeficient pořadové korelace se používají, pokud data nepocházejí z normálního rozdělení. Jsou tedy robustní v případě odlehlých pozorování. Spearmanův korelační koeficient se zaměřuje na pořadí jednotlivých měření obou proměnných, takže zachycuje rostoucí a klesající trend. Naproti tomu Kendallův korelační koeficient pořadové korelace se snaží měřit sílu vztahu dvou proměnných a statistiku založil na inverzích v pořadí. Má jinou interpretaci než Spearmanovo koeficient, protože se zabývá kladnou a zápornou asociací dvou konkrétních realizací proměnných (konkordací a diskordancí). Více o těchto koeficientech viz [23] a [37].

Parciální korelační koeficient

Parciální korelační koeficient měří sílu vztahu mezi proměnnými X a Y , je-li vyloučena ovlivňující proměnná Z . Vzorec lze nalézt v literatuře viz [23], [37], [7]. Tyto koeficienty opět mohou být různých typů.

Je možno testovat, zda jsou korelační koeficienty významné či nikoliv. Pro jednotlivé typy korelačních koeficientů lze testovat nulovou hypotézu o tom, zda jsou dvě proměnné navzájem nezávislé. Statistika má pro korelační koeficienty Studentovo rozdělení s $n - 2$ stupni volnosti. Jedná-li se o Kendallovu statistiku, používá se k rozhodování o nulové hypotéze z-statistika a kritické hodnoty normovaného normálního rozdělení. V případě parciálních korelačních koeficientů jsou stupně volnosti definovány jako $n - 3$, kde n je v tomto případě počet trojic dat ve výběru. Konkrétní popisy jsou k dispozici v literatuře viz [37]. Koeficient značí svou významnost, je-li zamítnuta nulová hypotéza na hladině významnosti α . Veličiny lze pak považovat za lineárně závislé.

3.3.3 Hodnocení průměru a rozptylu

Jsou-li data porovnávána, je důležitou vlastností střední hodnota a rozptyl. Proto tato část bude věnována testům, které hodnotí rozptyl a průměr. Porovnávat lze jednu proměnnou s teoretickou hodnotou, dvě proměnné mezi sebou a nebo více skupin měření proměnných mezi sebou. Je tedy rozlišováno mezi testy jednovýběrovými, dvouvýběrovými a vícevýběrovými. Nejprve budou představeny testy hodnotící průměr v jednom a ve dvou výběrech. Porovnáním více než dvou průměrů se bude zabývat kapitola 3.3.4.

T-test

K základnímu porovnání dvou populací je vhodné porovnávat jejich průměry. K tomu slouží t-test. Tato část krátce shrne problematiku jednoho a dvou výběrů.

Porovnávání průměru v **jednom výběru** se používá, pokud se práce zabývá otázkou, zda se liší výběrový průměr μ od teoretické hodnoty μ_0 při náhodném výběru dané populace. Testuje nulovou hypotézu $H_0 : \mu = \mu_0$. Jedním z příkladů může být například linka vyrábějící pneumatiky. Náhodně se vyberou některé z nich a testují se, zda průměr pneumatiky je stejný jako požadovaná hodnota průměru μ_0 .

Dvouvýběrový t-test slouží k porovnání průměrů dvou kategorií jedné proměnné (μ_1, μ_2). Typickým příkladem je rozdíl mezi pohlavím. Toto je

ukázka, kdy lze porovnávat dvě nezávislé výběry. Lze porovnávat i průměry ve dvou závislých výběrech (párová měření) například jsou-li sledovány účinky léků před a po aplikaci pro každého jedince. Hypotézy u jednostranného t-testu vypadají následovně $H_0 : \mu_1 = \mu_2, H_1 : \mu_1 > \mu_2$ ($\mu_1 < \mu_2$), u dvoustranného t-testu $H_0 : \mu_1 = \mu_2, H_1 : \mu_1 \neq \mu_2$. Je-li testovací statistika $t = \frac{\bar{x}_1 - \bar{x}_2}{s(\bar{x}_1 - \bar{x}_2)}$ větší než kritická hodnota t-testu t_α , $t_\alpha : P(t > t_\alpha) = \alpha$, je zamítnuta nulová hypotéza na hladině významnosti α . Dvoustranný interval spolehlivosti pro hladinu spolehlivosti $1 - \alpha$ má tvar $(\mu_1 - \mu_2) \in (\bar{x}_1 - \bar{x}_2) \pm t_{\frac{\alpha}{2}} s(\bar{x}_1 - \bar{x}_2)$.

Jsou-li k dispozici dva nezávislé výběry, je rozlišováno, zda rozptyly (s_1^2, s_2^2) jsou stejné či rozdílné. K tomu lze použít test významnosti rozdílu dvou rozptylů (F-test). Je-li počet pozorování v daných skupinách dostatečně velký ($n > 30$) a jsou-li známy rozptyly, používá se

$$s(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2(x)}{n_1} + \frac{s_2^2(x)}{n_2}}. \quad (3.10)$$

Vzhledem k tomu, že rozptyly jsou známy, lze je dosadit místo výběrových rozptylů a použijí se tabulky kritických hodnot pro standardizované normální rozdělení. Tento test je nazýván z-test průměrů. Postup lze provést i v případě nestejnosti rozptylů například jsou-li porovnávány různé populace. Test, který porovnává dva nestejně rozptyly je označován jako **Welchův t-test**.

Jsou-li rozptyly shodné v obou populacích, počet stupňů volnosti je $n_1 + n_2 - 2$ a směrodatná chyba je

$$s(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{(n_1 - 1) \cdot s_1^2(x) + (n_2 - 1) \cdot s_2^2(x)}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}. \quad (3.11)$$

Hodnocení rozptylu hraje roli při ověřování platnosti předpokladu u dvouvýběrového t-testu s nezávislými výběry. Je porovnáván rozptyl proti konkrétní hodnotě, dva rozptyly proti sobě a nebo rozptyly více skupin mezi sebou. V následující části bude krátce popsána problematika dvou a více výběrů. Hodnocení rozptylu v jednom výběru lze nalézt v literatuře viz [23].

F-test

F-test porovnává rozptyly ve dvou výběrech. Mají-li dva výběry s počty pozorování n_1, n_2 , rozptyly $s_1^2(x), s_2^2(x)$ normální rozdělení $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$, je testována nulová hypotéza $H_0 : \sigma_1^2 = \sigma_2^2$ oproti alternativní hypotéze

$H_1 : \sigma_1^2 \neq \sigma_2^2$. F-test je poměr mezi jednotlivými rozptyly. Vychází z χ^2 rozdělení, jež se používá při zkoumání variability dat. Poměr dvou nezávislých χ^2 rozdělení má Fisherovo-Snedecorovo rozdělení se stupni volnosti $(n_1 - 1, n_2 - 2)$. Testovací statistiku $F = \frac{s_1^2(x)}{s_2^2(x)}$ je porovnávána s kvantilem F-rozdělení se stupni volnosti n_1, n_2 a s hladinou významnosti $1 - \frac{\alpha}{2}$. Nulová hypotéza je zamítnuta, jestliže testovací statistika je větší než kvantil F-rozdělení. [23]

Levenův test

Další test, který zkoumá shodnost či rozdílnost více rozptylů, je Levenův test shody rozptylů. Nulová hypotéza je definována jako $H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2$, kde n je počet skupin. Tento test se používá pro ověření předpokladu o shodě rozptylů (homoskedasticitě) v daných výběrech před provedením analýzy rozptylu (ANOVA). Ovšem je možno použít i jiný test, například Bartlettův nebo Brownův-Forsytheův test.

Levenův test vychází z dvouvýběrového t-testu nezávislých skupin, ovšem místo daného pozorování x_i se použije absolutní odchylka (či kvadratická odchylka) pozorování od skupinového průměru, jež charakterizuje rozptýlenost pozorování. [23]

Brownův-Forsytheův test nahrazuje průměr mediánem, protože medián vykazuje lepší robustnost vůči nestandardním datům (např. nesplňují předpoklad normality). Bartlettův test je citlivější na porušení předpokladu normality proměnné než předešlé dva testy. Při nesplnění podmínky se může zvýšit pravděpodobnost chyby I. druhu.

V případě, že je porušen jeden z předpokladů dvouvýběrového t-testu (normalita, shodnost rozptylů), používá se Wilcoxonův pořadový test, který je shrnut níže.

Wilcoxonův pořadový test

Tento test (angl. *Wilcox rank sum test*) je řazen mezi neparametrické testy. Jedná se o neparametrickou variantu t-testu, používá se pro porovnání dvou nezávislých výběrů. Při výpočtu testové statistiky jsou seřazeny hodnoty z obou výběrů a jsou jim přiřazena pořadová čísla. Součet pořadových čísel vstupuje do výpočtu testového kritéria. Konkrétní popis testu lze najít v literatuře viz [37] a [13].

3.3.4 Analýza rozptylu

Analýza rozptylu (angl. *analysis of variance*, zkr. ANOVA) je metoda, která zkoumá vztahy mezi kategorickými proměnnými (skupinami) a závislou spojitou proměnnou. Zkoumá, zda jsou rozdíly mezi skupinami statisticky významné. Často se poukazuje na rozdíly mezi pohlavím, leváky a praváky či věkovými skupinami. Pokud se porovnávají průměry dvou populací, používají se dvouvýběrové testy zmíněné výše. Ovšem pokud jsou porovnávány více než dva průměry, používá se metodika ANOVA.

Nezávisle proměnná (faktor) může nabývat až k hodnot (úrovní, kategorií). Je rozlišováno mezi jedno faktorovou a více faktorovou analýzou rozptylu. Zkoumání jednoho faktoru na závisle proměnnou je nejjednodušším případem analýzy rozptylu. Ovšem je-li faktorů více, jsou zkoumány hlavní a interakční efekty. [23]

V analýze rozptylu se používá F statistika rozdílu průměrů daných skupin. Jedná se o poměr váženého rozptylu průměrů hodnot ve skupinách a rozptylu hodnot v dané skupině. Testuje se, zda průměry skupin se od sebe významně liší či skupiny nemají vliv na závisle proměnnou a tudíž jsou stejné. Pokud je F-statistika větší než kritická hodnota, zamítá se nulová hypotéza o shodnosti průměrů.

Před samotným testováním je ale třeba ověřit předpoklady ANOVA testu. Data musí pocházet z normálního rozdělení (normalitu lze otestovat buďto testy normality nebo pomocí vykreslení histogramů), výběry jsou nezávislé (lze usoudit podle metody sběru dat) a rozptyl všech výběrů je stejný (lze použít testy o shodě rozptylů). K tomuto účelu se pro $k = 2$ využívá se F-test, pro větší počet kategorií ($k > 2$) poté Levenův test. Pokud není splněn buďto předpoklad normality nebo shodnosti rozptylů v jednotlivých skupinách, používá se Kruskalův-Wallisův test.

Jednofaktorová analýza rozptylu

Jednofaktorová analýza rozptylu (angl. *one-way analysis of variance*) je rozšíření dvouvýběrového t-testu pro více nezávislých výběrů. Zkoumá tedy rozdíly středních hodnot daných skupin. Testuje nulovou hypotézu o shodnosti průměrů, tzn. o tom, že neexistuje rozdíl mezi skupinami $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ oproti alternativní hypotéze, že existuje alespoň jedna dvojice i, j , pro kterou platí $\mu_i \neq \mu_j$ a tudíž faktor ovlivňuje data. Podstatou testovací F-statistiky je to, že celková variabilita dat je vyjádřena pomocí celkového součtu čtverců SS_T , jež rozlišuje rozdíly mezi skupinami SS_A a uvnitř skupin SS_E . Ze součtů čtverců rozdílu lze podělením stupňů volnosti ($k - 1, n - k, n - 1$) získat průměrné čtverce MS_A, MS_E . [17]

Přehled zmíněných statistik je vidět níže.

$$SS_A = \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2 \quad (3.12)$$

$$MS_A = \frac{SS_A}{k-1} \quad (3.13)$$

$$SS_E = \sum_{i=1}^k \sum_{j=1}^{n_j} (x_{ij} - \bar{x}_i)^2 \quad (3.14)$$

$$MS_E = \frac{SS_E}{n-k} \quad (3.15)$$

$$SS_T = SS_E + SS_A = \sum_{i=1}^k \sum_{j=1}^{n_j} (x_{ij} - \bar{x})^2 \quad (3.16)$$

Výsledky výpočtů se zapisují do tabulky, jejíž schéma lze vidět v tab. 3.1. Může být spočítána i hodnota η^2 (viz [36]), jež je podílem součtů čtverců rozdílu celkové variability SS_T a variability mezi skupinami SS_A . Vyjadřuje jaký podíl rozptylu je vysvětlen nezávislou proměnnou.

Tabulka 3.1: Schéma tabulky ANOVA

Zdroj variability	Součet čtverců	Stupně volnosti	Odhad rozptylu	F-statistika
faktor A (mezi skupinami)	SS_A	$k-1$	MS_A	F
reziduální (uvnitř skupin)	SS_E	$n-k$	MS_E	
celkový	SS_T	$n-1$		

Kruskalův-Wallisův test

Tento test je neparametrickou analogií jednofaktorové analýzy rozptylu. Jedinými předpoklady k tomuto testu jsou nezávislé pozorování a spojitě rozdělení. Test porovnává, zda měření ve skupinách mají stejné mediány tj. $H_0 : \tilde{x}_1 = \tilde{x}_2 = \dots = \tilde{x}_k$. Testová statistika 3.17 využívá pořadí hodnot a pořadové statistiky. Pro velké hodnoty n_i se statistika blíží přibližně χ^2 rozdělení s $k-1$ stupni volnosti. [23]

$$KW = \frac{12}{n \cdot (n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3 \cdot (n+1), \text{ kde} \quad (3.17)$$

R_i ... součet pořadí v i -té skupině

V případě, že je zamítnuta hypotéza o shodě středních hodnot či mediánů, znamená to, že faktor A (mezi skupinami) ovlivňuje danou veličinu. Jestliže má být zjištěno, které ze skupin se liší, provádí se **mnohonásobné porovnávání**. Lze použít například párové srovnání mezi skupinami pomocí Tukey metody založené na t-testu (angl. *Tukey's Honestly Significant Difference test*, viz [5]) či v případě neparametrické obdoby analýzy rozptylu párový Wilcoxonův pořadový test (angl. *Pairwise Wilcoxon Rank Sum Test*). Tento test je modifikací Wilcoxonova pořadového testu, jež porovnává skupiny mezi sebou. Vzhledem k tomu, že se porovnává každá skupina s každou, násobí se tak i chyba I. druhu. Proto tento test uvažuje korekci pro výpočet p-hodnot. V našem případě je použita Holmova korekce, aby se chyba I. druhu nezvětšovala.

3.3.5 Analýza hlavních komponent a shlukování

Analýza hlavních komponent (angl. *Principal Component Analysis*, zkr. PCA) se používá k redukci dimenze dat pomocí tzv. hlavních komponent, jež popisují vazby a vztahy mezi daty. Pomocí komponent se lze dozvědět, zda nejsou v datech odlehlá pozorování či zda není narušena vazba mezi proměnnými (nehomogenita skupin, nestejnost rozptylů, nerovnoměrné rozdělení hodnot, nezávislost pozorování, apod.). [22]

Hlavní komponenty vznikají ortogonální lineární transformací vícerozměrných původních dat do množiny dat s nižší dimenzí. Vznikají nové proměnné, které jsou nekorelované (mají-li původní data normální rozdělení, tak i nezávislé), takže každá z nich měří jinou vlastnost dat. Hlavní komponenty jsou seřazeny podle velikosti rozptylu. Zajímavé jsou ty komponenty, které vysvětlují nejvíce rozptylu původních dat (zbytek se zanedbává). První komponenta vysvětluje co nejvíce z celkové variability dat (součet rozptylů zkoumaných proměnných), další komponenta vysvětlí co nejvíce ze zbývající variability dat a je kolmá na první, atd. Vhodnou vizualizací může být Paretův graf, jež zobrazuje četnost procentuálních vysvětlení variability dat pro jednotlivé komponenty a kumulativní četnost těchto vysvětlovaných rozptylů. Z něj lze vidět, kolik komponent vysvětlí co nejvíce variability. Rozhoduje se, které hlavní komponenty jsou významné (je možné data popsat bez ztráty informace) a tedy tedy jakou dimenzi má smysl uvažovat při redukci dat. Většinou stačí, aby vybrané komponenty vysvětlily 80 % variability dat. [23]

První hlavní komponenta je tvořena lineární kombinací vektoru a proměnných. Součet složek vektorů musí být roven 1, je tedy normovaný a je zajištěno co největšího rozptylu této komponenty. Proměnné jsou na začátku

znormovány. Další komponenty jsou opět kombinace normovaného vektoru a proměnných, jež se snaží o co možná největší rozptyl. Při hledání hlavních komponent se vychází z kovarianční matice, jež je symetrická a pozitivně semidefinitní. Pro získání vlastních čísel (nezáporných) a jim odpovídajících ortonormálních vlastních vektorů (jsou na sebe kolmé, délka vektoru je rovna jedné, kovariance je nulová) je použit singulární rozklad (SVD) symetrické kovarianční matice. Hlavní komponenty jsou pak dány kombinací vlastních vektorů a původních proměnných, jež mají nulovou střední hodnotu. Geometricky se jedná o projekci a natočení vícerozměrných dat do nejčastěji dvourozměrného prostoru či jednorozměrného prostoru. Vícerozměrná data tvoří shluk bodů připomínající elipsoid či hyperkouli. Pokud se jedná o projekci do dvourozměrného prostoru, vykreslí se pouze dvě hlavní komponenty, jež se natočí podle soustavy souřadnic. Z důsledku ortonormality jsou komponenty na sebe kolmé a ztrácí tak určitou informaci o původních datech.

Hlavní komponenty lze vykreslit v závislosti na kategoriích. Využívá se tak poznatků z problematiky **shlukování** (angl. *clustering*). Shluky jsou skupiny (objekty), v nichž jsou si prvky podobné (stejnorodé), ale mimo ně rozdílné. Existuje několik metod, jak tyto shluky najít. V případě hierarchických metod se algoritmy snaží eliminovat počet shluků na minimum (na začátku tvoří každý bod shluk, tvoří hierarchickou stromovou strukturu). V případě nehierarchického shlukování je třeba určit počet shluků k (např. jako počet kategorií, jež je třeba mít). Používá se metoda K průměrů (angl. *K-means clustering*), iterativně optimalizuje vzdálenost bodů od středu shluku (centroidu) tak, aby euklidovská vzdálenost byla od daného centroidu co nejmenší (na začátku je rozdělena množina dat do k shluků). V každém kroku algoritmus vypočítá nové reprezentanty centroidů a přepočítá vzdálenosti od všech bodů. Body přiřazuje k těm shlukům, jež má menší vzdálenost od centroidu. Algoritmus se opakuje do té doby, dokud se shluky mění. [37]

4 Programové řešení

Značná část práce byla věnována tvorbě nástroje pro statistické zpracování dat, který lze použít na libovolná data ve formátu *.xls* a *.xlsx* obsahující kategorické i numerické proměnné. Nemusí se tedy data týkat pouze antropometrie rukou. Tato část práce je věnována popisu toho, jak byla aplikace vypracována, které hlavní problémy musely být řešeny, jaká jsou omezení aplikace a také to, jak byla provedena kontrola naimplementovaných algoritmů.

4.1 Vypracování aplikace

Statistické zpracování dat bylo naprogramováno v programovacím jazyce R za použití knihovny *Shiny* jako webová aplikace. Díky platformě *shinyapps.io* (angl. *Platform as a Service*, zkr. PaaS) je možné sdílet vytvořené Shiny aplikace na web. Vzhledem k tomu, že tato platforma je samoobslužná, je tento způsob jednodušší než v případě konfigurace vlastního serveru a další náležitě síťové konfigurace v případě, že by aplikace měla být vystavena na internet. [40]

Před započítím práce bylo třeba zjistit jaké webové technologie by se daly použít. Cílem bylo vytvořit aplikaci se silným důrazem na statistické zpracování dat. A toto byl hlavní důvod, proč byly vyřazeny technologie jako ASP.NET (C#), JSP (Java) a PHP. Dala by se ještě použít knihovna Django pro programovací jazyk Python. Sice by výše zmíněné technologie nabízely větší volnost při tvorbě webové aplikace, té však nebylo potřeba, protože je práce cílena na tvorbu aplikace s úzkým zaměřením. I nejen proto byla vybrána technologie Shiny postavená na volně dostupném programovacím jazyce R. Vzhledem k tomu, že jazyk R se vyvíjí komunitou už 26 let, nabízí mnoho knihoven i optimální výkon pro statistické zpracování dat, což například jazyk C# ve volně dostupné formě nenabízí.

Shiny je určena pro snadný vývoj malých až středně velkých interaktivních webových aplikací zaměřených na analýzu dat a jejich grafické znázornění. Skládá se z knihovny *Shiny* a webového serveru. Knihovna poskytuje příkazy pro tvorbu webových aplikací a server tyto příkazy zpracovává na základě vstupu od uživatele. Nejen knihovna, ale i webový server jsou volně dostupné. Pro vývoj aplikace lze použít libovolné vývojové prostředí podporující jazyk R. V našem případě bylo použito prostředí RStudio, které je také dostupné jako open-source. Webový server existuje ve třech varian-

tách. První variantou je, že je server součástí vývojového prostředí RStudio. Tato varianta je doporučena pro vývoj, testování a lazení (angl. debugování) aplikace. Druhá varianta je dostupná pouze pro server s operačním systémem Linux a jedná se o samostatný webový server. Tato varianta je vhodná pro nasazení aplikace v produkčním prostředí. Poslední variantou je využití cloudového řešení *shinyapps.io* poskytované společností RStudio. Tato varianta umožňuje Shiny aplikace snadno vystavit na internetu.

Shinyapps.io poskytuje možnost bezplatného běhu až 5 aplikací nasazených pod jedním účtem. Dalším omezením bezplatné varianty je maximálně 25 aktivních hodin měsíčně. Jako aktivní čas se započítává čas, kdy alespoň jeden klient (webový prohlížeč) aplikaci využívá (má ji načtenou). Právě této možnosti bylo využito. Proto se také může stát, že přidělený čas bude vyčerpán a aplikace se stane na platformě shinyapps.io nedostupnou. K dispozici jsou další varianty navyšující počet aktivních hodin za poplatek.

Shiny může být použito pro tvorbu samostatných webových stránek, vložení do dokumentů technologie R Markdown (obdoba technologie Jupyter pro Python, jedná se o interaktivní dokumenty) nebo tvorbu jednoduchých nástěnek (angl. dashboardů). I bez konkrétního nastavení Shiny zajistí defaultní CSS téma (angl. CSS theme) a JavaScriptové akce na straně klienta. Ty však mohou být nahrazeny vlastními. V práci bylo zvoleno grafické CSS téma *flatly* a byla rozšířena možnost ovládání webové stránky na straně klienta pomocí knihovny *ShinyJS*.

4.2 Postup při vytvoření aplikace

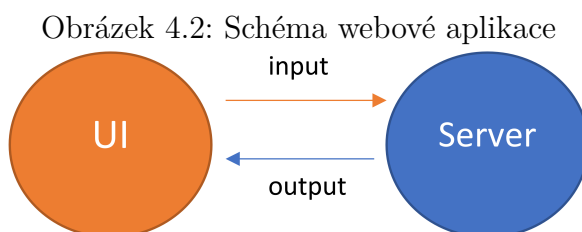
Aplikace byla navržena tak, aby se po nahrání libovolného souboru formátu *.xlsx* či *.xls* provedla analýza dat. Aplikace je vizuálně rozvržena do tří částí - hlavní nabídky (menu), panelu možností (nalevo) a výstupního panelu. Náhled na přehledovou tabulku načtených dat lze vidět na obr. 4.1.

Obrázek 4.1: Náhled webové aplikace

Population Group	Count	Min	Max	Mean	Median	Standard Deviation	Variance	Variance Coefficient	Standard Error
20-29	109	151.00	190.00	175.00	175.00	8.74	76.45	4.98	0.28
30-39	58	160.00	191.00	177.00	177.00	7.75	60.06	4.36	0.37
40-49	97	160.00	191.00	177.00	177.00	6.87	47.15	3.87	0.27
50-59	88	151.00	190.00	175.00	175.00	6.89	47.46	3.93	0.28
60-69	43	160.00	197.00	174.00	174.00	6.80	77.45	5.06	0.45
70-79	27	157.00	180.00	174.00	174.00	7.43	55.15	4.27	0.46
All	432	155.00	197.00	175.00	175.00	7.80	60.84	4.43	0.13

Sex	Count	Min	Max	Mean	Median	Standard Deviation	Variance	Variance Coefficient	Standard Error
Male	109	90.00	112.00	100.00	100.00	5.05	25.51	1.05	0.22
Female	58	86.00	109.00	97.00	97.00	4.65	21.61	4.60	0.29
All	97	89.00	110.00	100.00	100.00	4.88	23.52	4.64	0.22
Male	88	86.00	103.00	95.68	95.00	4.30	17.41	4.21	0.22
Female	43	92.00	112.00	99.51	98.00	4.73	22.35	4.75	0.33
Male	27	89.00	110.00	99.43	100.00	5.05	25.47	5.00	0.37
All	432	86.00	113.00	96.29	96.00	6.72	45.13	4.71	0.30

Struktura vstupního bodu aplikace je rozdělena na klientskou (definice uživatelského rozhraní) a serverovou část (chování aplikace). Schéma je vidět na obr.4.2. Serverová část přijímá požadavky z uživatelského rozhraní (webový prohlížeč) a zpracovává je. Následně jsou výsledky odeslány zpět do uživatelského rozhraní. Jedná se o výstupy z algoritmů. Znamená to, že všechny výpočetní operace jsou prováděny na serveru. Na klientské straně si sama technologie Shiny řeší běh jednoduchých JavaScriptů starající se o komunikaci a události v uživatelském rozhraní (reaktivitu). Popis a práce s aplikací je k nalezení v uživatelské příručce viz B.3.



Prvním úkolem bylo načtení dat. Původní načtení dat pomocí funkce *read.xlsx()* z knihovny *openxlsx* načítalo pouze soubory s příponou *.xlsx* a navíc po nahrání souboru byly mezery v hlavičce (názvy proměnných) nahrazeny tečkou. Proto bylo rozhodnuto použít funkci *read_excel()* z knihovny *readxl*. Aby bylo zaručeno toho, že uživatel vybere pouze soubory daného typu, bylo třeba tuto restrikcí definovat v kódu aplikace. To bylo zajištěno specifikací přípon a MIME typů ¹ odpovídajícím výše uvedeným formátům dat. Tyto informace byly předány uživatelskému prvku Shiny *fileInput()*.

Listing 4.1: Definice ovládacího prvku pro výběr souboru

```

fileInput("data",
  "Load Excel spreadsheet",
  multiple = FALSE,
  accept = c("application/vnd.ms-excel",
    "application/vnd.openxmlformats-officedocument.
spreadsheetml.sheet",
    ".xls", ".xlsx"))
  
```

¹MIME typ je obdoba přípony u názvu souboru uloženého na disku. Definuje jakým způsobem má prohlížeč a server nakládat s přenašenými daty. Například pro starší soubory aplikace Excel s příponou *.xls* je MIME typ *application/vnd.ms-excel*.

V aplikaci lze načíst dva datové soubory - data s naměřenými hodnotami a data s informacemi pro ostatní populace. V případě porovnání s ostatními populacemi bylo vhodné, aby uživatel mohl v případě potřeby editovat tabulku podobně jako v Excelu. Pro tento účel byly použity možnosti knihovny *rhandsontable*, jež nabízí *handsontable* widget (složitější ovládací prvek). Aby upravená data mohla být použita, musela být převedena do objektu, s kterým R dokáže pracovat pomocí funkce *hot_to_r()*.

Listing 4.2: Inicializace a použití *handsontable*

```
#pri stisku tlacitka , nacti data
  observeEvent(input$loadDataCompare ,
  #ted se data nactaji az po stisku tlacitka loadData
    {
output$dataCompContents <-
renderRHandsontable({ rhandsontable(myDataComparison())})
#zobrazeni nactene tabulky
... })

#chci ulozit zmenene polozky pri stisku tlacitka SAVE
  observeEvent(input$saveBtn , {
#Convert handsontable data to R object
  comparisonValues$comparisonData <-
  hot_to_r(input$dataCompContents)})
```

Vzhledem k tomu, že v serverové části bylo zamýšleno načíst data pouze jednou (ne aby při každém použití proměnné se načítali znovu data ze souboru) a ty si předávat mezi vnořenými funkcemi, byla použita funkce *reactive()*. Tato funkce aktualizuje svou hodnotu pouze za stanovených podmínek, v naší práci pouze jednou při prvním použití stisku tlačítka *Load data*. Funkce uloží hodnotu do mezipaměti, ze které ji ostatní části kódu čtou. Pokud by se podmínka změnila (např. načtením nových dat), výraz je opět proveden a data v mezipaměti jsou aktualizována. V serverové části je také funkce *observeEvent()*, jež reaguje na událost a tím spustí určený kód (reaguje na *triggery*). V našem případě se jedná o stisk tlačítka, jež zapříčiní spuštění příslušné sady příkazů.

Listing 4.3: Použití funkce *reactive*

```
myData <- reactive({
  if (is.null(input$data)) { return(NULL) }
```



```

#neukonci svouj beh pri nevybranem souboru
predzpracovani_2(as.data.frame(read_excel(
  path= input$data$datapath ,
  sheet = input$sheetxls ,
  na = c("NA", " "))))
})

```

Při načítání souboru bylo zároveň zamýšleno, aby byly vymazány řádky, jež obsahují neplatnou hodnotu ("", 9999) a odlehlá pozorování. Zároveň bylo žádané, aby uživatel měl informaci o tom, které řádky (id probandů, id odlehlých hodnot) byly smazány, aby měl možnost datový soubor případně upravit. Algoritmus je k nahlédnutí v příloze viz B.1.

Při základních charakteristikách je třeba, aby se pro kategorické a numerické proměnné provedl jiný způsob zpracování dat a zobrazení konkrétních grafů a tabulek. K tomu byla využita knihovna *shinyjs*, pomocí které lze volat JavaScript zpracovávaný prohlížečem ze serveru. Tato knihovna byla vybrána, aby bylo docíleno lepší interaktivity (a tudíž lepší uživatelské zkušenosti) naší aplikace. Z knihovny *shinyjs* bylo využito příkazů pro skrytí a zobrazení konkrétních tabulek či grafů v závislosti na nastavení určených uživatelem. Tato funkcionality není v Shiny standardně podporována. Protože knihovna *shinyjs* poskytuje funkce *show()* a *hide()* jsou tyto funkce s celkem běžnými názvy "překryty" stejnojmennými funkcemi z jiných knihoven, jež jsou použity. Z tohoto důvodu bylo třeba použít explicitní vyjádření knihovny a její funkce za pomoci dvojité dvojtečky (např. *shinyjs :: show("density")*, *shinyjs :: hide("barPlot")*). Před použitím těchto příkazů je nutno inicializovat samotnou knihovnu vložení neviditelného uživatelského prvku *useShinyjs()*.

Dále bylo třeba zajistit dynamické vytváření grafů (uživatelských prvků pro vykreslení grafu). Nejdříve bylo nutné vygenerovat jednoznačné názvy pro požadované grafy, grafy vyrenderovat a uložit do seznamu obsahující uživatelské prvky jako nové objekty. Následně byly tyto předem vyrenderované grafy určeny k vykreslení v konkrétním ovládacím prvku.

Listing 4.4: Příklad přípravy dynamického množství grafů

```

#nejdriv si musime jednotlivy grafy vyrenderovat a ulozit
for (i in 1:length(nameVariables)) {
  local({
#local je zde nutne, aby kazdy graf mel sve unikatni cislo
    aktualni_i <- i #tohle je taky nutne
    plotName <- paste("boxplot", aktualni_i, sep=" ")
  })
}

```

```

#vytvorime si jmeno plotu
output[[plotName]] <- renderPlot({
#vyrenderujeme ho a ulozime do outputu

tabulka<-data.frame(cond=group ,
  xvar=X[,aktualni_i])

ggplot(data=tabulka , aes(x=tabulka$cond ,
  y=tabulka$xvar ,
  fill=tabulka$cond)) +
  geom_boxplot()+
  ggtitle(paste("Boxplot of ",
    nameVariables[aktualni_i] ,
    sep=" "))+
  xlab("Category")+
  ylab("Values")+
  theme(plot.title=element_text(lineheight=.8,
    face="bold"))+

  stat_summary(fun.y=mean,
    colour="darkred" ,
    geom="point" ,
    size=3)+
  scale_fill_discrete(name="Category"))
})}

```

Dalším krokem bylo zobrazení grafů v uživatelském prvku *uiOutput* příkazem *renderUI*. V tomto příkazu všechny předem vyrenderované grafy byly připraveny pro výstup pomocí příkazu *plotOutput* do seznamu a následně je pomocí příkazu *do.call* vygenerován konkrétní HTML kód jako výstup pro *renderUI*.

Listing 4.5: Příklad dynamického množství grafů

```

output$boxplots <- renderUI({
  dynamicOutputList <- lapply(1:length(nameVariables) ,
  #sestavime si seznam vsech plotu , ktere jsme si
  #pred-renderovali
  function(i){ #anonymni funkce
    name <- paste("boxplot" , i , sep=" ")
    plotOutput(name)
    # zde si pred-renderovany graf

```

```

    # vykreslime do seznamu
  })
do.call(tagList, dynamicOutputList)
#a timto prikazem ho posleme do renderUI
})

```

Nemalé úsilí bylo věnováno tvorbě tabulek a grafů, jež jsou celou aplikací propleteny. Některé popisky k jednotlivým tabulkám a grafům byly vytvořeny pomocí HTML tagů. K tomu byla využita funkce *tags()*. Příklady některých vybraných částí jsou uvedeny v přílohách viz B.2 a B.3.

4.2.1 Aktivace účtu na shinyapps.io

Vzhledem k tomu, že bylo rozhodnuto, aby aplikace byla hostována na cloudu, bylo třeba vytvořit účet na *shinyapps.io*. K tomu bylo potřeba mít vývojové prostředí RStudio ² a nainstalovat nejnovější verzi knihovny *rsconnect*, která nasadí aplikaci na *shinyapps.io* službu. Tato knihovna je nyní hostována na CRANu (angl. *Comprehensive R Archive Network* ³). Dále bylo třeba si založit účet na *shinyapps.io* a provést konfiguraci knihovny *rsconnect*, aby používal zadaný účet. Automaticky byl vygenerován token s heslem, který byl zadán do konzole RStudia. [28]

Listing 4.6: Autorizace vytvořeného účtu na shinyapps.io

```

rsconnect::setAccountInfo(name='datatesting',
                          token='.....',
                          secret='<SECRET>')

```

4.2.2 Nasazení aplikace

Po úspěšné autorizaci stačilo už jen nasadit aplikaci na web pomocí příkazu *deployApp()*. Nicméně v prostředí RStudio lze publikovat aplikaci přes stisk tlačítka *Publish*. Více informací lze nalézt na stránkách *shinyapps.io* viz [28].

Listing 4.7: Nasazení aplikace na web

```

library(rsconnect)
rsconnect::deployApp('<cesta_k~aplikaci>')

```

²Instalace běhového a vývojového prostředí je popsána v příloze viz B.1.

³CRAN je síť FTP a webových serverů ukládající aktuální verze kódu a dokumentace pro R.

4.3 Omezení a možné problémy s vytvořenou aplikací

V aplikaci je podporován pouze formát *.xls* nebo formát *.xlsx*. Jiné formáty (např. *.csv*) nejsou podporovány. Stane-li se, že uživatel zadá název listu souboru, který neexistuje, aplikace ukončí svůj běh. U porovnání s ostatními populacemi musí načítaný soubor splňovat danou strukturu. Prvním sloupcem je pojmenování národnosti/státu, druhým je pohlaví (M/F, jiná kategorizace není možná), třetím je počet pozorování, další jsou pro rozměry (jsou uvedeny 2 informace - průměr a směrodatnou odchylku v tomto pořadí pro každý rozměr).

Načtení souboru do aplikace je rychlé, nicméně zobrazení velké tabulky a vykreslení grafů může trvat déle (řádově jednotky vteřin). Je třeba si uvědomit, že nejsou-li data nahraná, nelze je vybírat. Dále pak odlehlá pozorování a neplatné hodnoty jsou odstraňovány automaticky (uživatel nemá volby).

Při nevhodném výběru proměnných nebo kategorií (např. u parciálních korelačních koeficientů nebudou vybrány 3 odlišné proměnné, bude vybrána skupina s více než 10 kategoriemi) se určité tabulky buďto vůbec nezobrazí a vypíšou chybovou hlášku nebo graf nebude přehledný. V pokročilých statistikách (hlavní nabídka *Advanced Statistics*) je třeba zaškrtnout více než jednu proměnnou. V opačném případě aplikace ukončí svůj běh. Rovněž je třeba u porovnání více populací (hlavní nabídka *Comparison*) zaškrtnout volbu ženy nebo muže, aby se provedly výpočty.

Dále během vývoje aplikace bylo zjištěno, že v případě funkcí *cor.test()* a *ks.test()* je zde podmínka na tzv. *ties*. Jedná se o shodné hodnoty, jež se několikrát opakují v datech konkrétní proměnné. Proto se v konzoli RStudia vypisovala upozornění, že p-hodnota může být nepřesná. V našem případě byly p-hodnoty testů, které jsou menší než 0,001 a větší než 0, zapsány symbolem < 0.001 . Pokud p-hodnota je 0, znamená to, že nemohla být vypočítána.

Ne všechny znaky s interpunkcí v názvu proměnné jsou podporovány. Proto je vhodné proměnné pojmenovat bez interpunkce. Také je vhodné, aby soubor obsahoval nejen numerické proměnné, ale i kategorické proměnné, protože určité tabulky a grafy jsou vykresleny v závislosti na vybrané kategorii.

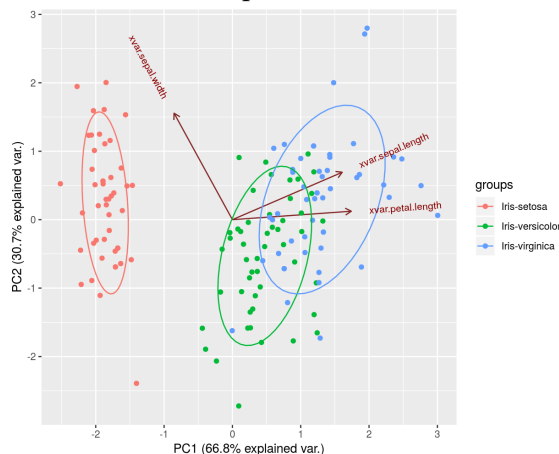
Jak bylo uvedeno výše, platforma Shinyapps.io poskytuje možnost bezplatného využití 5 spuštěných aplikací, jež běží v součtu 25 hodin měsíčně.

4.4 Testování

V této sekci je popsáno, jak byla ověřena kontrola napsaných algoritmů. Aplikaci je možno otestovat přímo na stránce <https://datatesting.shinyapps.io/final/> nebo lze aplikaci spustit na vlastním počítači pomocí postupů uvedeného v příloze B.2. Testován byl hlavně námi napsaný kód, kód použitých knihoven byl testován během vývoje jen v některých případech (spoléhá se na kvalitu veřejně dostupných, komunitou vyvíjených a používaných knihoven).

Aplikace byla testována s veřejně dostupným datasetem *iris setosa*, kde výstupy jednotlivých funkcí z aplikace byly porovnány s veřejně dostupnou analýzou těchto dat. Dataset obsahuje 5 proměnných (1 kategorickou o druzích kosatců, 4 numerické - rozměry kalichů a okvětních lístků). Celkem bylo nahráno 148 záznamů (2 záznamy s nejvyššími hodnotami šířky kalichu druhu *iris setosa* byly identifikovány jako odlehlá pozorování a tudíž vymazány). Mezi skupinami byly potvrzeny statisticky významné rozdíly (ANOVA). Metoda hlavních komponent redukovala dimenzi úlohy na 2. První dvě komponenty vysvětlují 95,7 % celkové variability původních dat. V grafu 4.3 lze usoudit, že lze pomocí délky a šířky okvětních lístků identifikovat druh kosatců (*iris setosa*, *versicolor* a *virginica*).

Obrázek 4.3: Metoda hlavních komponent - dataset *iris setosa*



Dále byl vytvořen dataset (`data_outliers.xlsx`) na zkoušku odlehlých pozorování, aby bylo zjištěno, zda je metoda MAD implementována správně. Dataset obsahuje hodnoty 1,2,3 (28 hodnot) a poté 13 a 100. Algoritmus detekoval místa, kde byly vysoké hodnoty a záznamy s čísly 13 a 100 odstranil.

Další dataset (`data_na.xls`) obsahoval neplatné hodnoty, kdy chybějící hodnoty, NA hodnoty nebo symboly s číslem 9999 byly odstraněny. Zároveň byla provedena kontrola, na práci se souborem ve formátu `.xls`.

Dále bylo otestováno chování programu za různých nastavení v uživatelském rozhraní. Rovněž i editování nahraného souboru v sekci *Comparison*. Jak bylo zmíněno v omezení, některé metody vyžadují výběr dvou a více proměnných, jinak aplikace nečekaně ukončí svůj běh.

Při načítání souborů je odezva závislá na množství dat, rychlostí spojení mezi prohlížečem a serverem a rychlostí samotného serveru. Doba od stisku tlačítka *Load data* až po zobrazení tabulky trvá řádově několik vteřin. Jak bylo zjištěno, většinu tohoto času zabírá renderování tabulek a grafů.

5 Analýza dat

Tato kapitola se věnuje výsledkům analýzy dat. Data o antropometrii rukou byla získána v rámci výzkumu, jež se zabýval vývojem krátké ruční zbraně s ohledem na ergonomii a antropometrii člověka, z Regionálního technologického institutu (RTI) na Západočeské univerzitě v Plzni v letech 2013-2015. Proto je tato část věnována analýze rozměrů rukou.

Pokud jsou analyzována jakákoliv data, mělo by být kontrolováno i to, jak byla získána. Proto je první část věnována sběru dat a tvorbě antropometrické databáze. Při práci s daty je ovšem kontrolováno i to, zda se v nich nevyskytují nějaké neúplné hodnoty nebo odlehlá pozorování, a proto jsou odstraněna. Dále je provedena podrobná analýza zkoumaných dat.

5.1 Sběr dat

Při tvorbě antropometrické databáze měření rukou českých probandů bylo třeba postupovat podle pravidel a zásad evropských norem, jež jsou technickými specifikacemi výrobků, služeb a postupů [2]. Konkrétně normy ČSN EN ISO 7250 - Základní rozměry lidského těla pro technologické projektování [3] a ČSN EN ISO 15535 - Všeobecné požadavky na zakládání antropometrických databází a dle přidružené normy [19] byly zvoleny z toho důvodu, aby se data dala porovnat s různými studiemi. V těchto normách jsou k nalezení potřebné informace, které by měly být dodrženy. Jedná se o uvedené rozměry, informace o probandech, formát databáze, obsah databáze, postup při statistickém zpracování, počet osob (měření), rozdělení skupin podle věku.

Norma ČSN EN ISO 7250 obsahuje rozměry, popisy rozměrů, měřicí metody i doporučená měřidla. Rozlišuje následující délky:

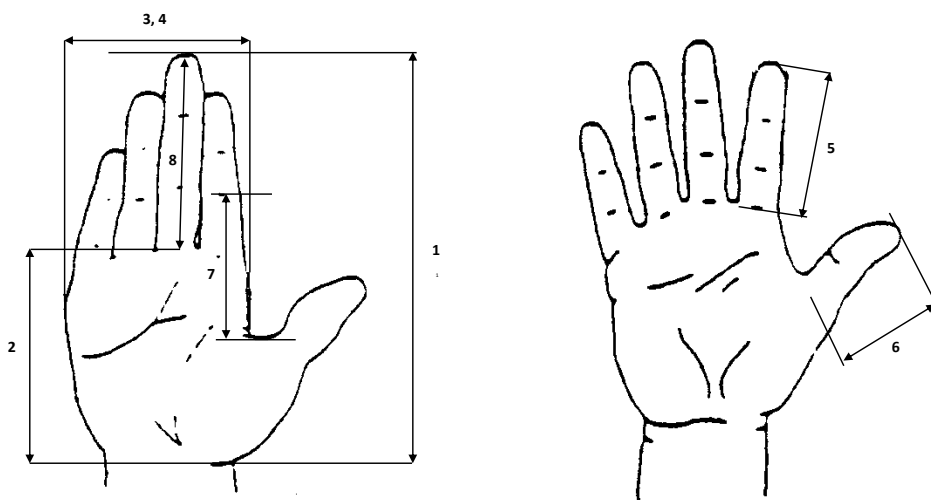
- dlaňová délka ruky, (dále jen délka ruky, 1)
- délka dlaně (2)
- šířka dlaně ve výši metacarpalia (3)
- délka ukazováku (5)

Další délky uvedené v naší práci jsou:

- obvod dlaně (4)
- délka palce (6)
- délka od kořene palce k prvnímu kloubu ukazováku (dále jen délka pro stisk spouště, 7)
- délka prostředníku (vypočtená z dlaňové délky ruky a délky dlaně, 8)

Všechny měřené délky jsou zakresleny v obr. 5.1. Délky byly měřeny digitálním posuvným měřítkem a krejčovským metrem s přesností na 1 mm.

Obrázek 5.1: Měřené rozměry ruky



Dalšími dvěma rozměry jsou tělesná výška a tělesná hmotnost probanda. Pro měření výšky byl použit antropometr a pro zjištění tělesné hmotnosti osobní váha. Ta byla zaznamenána s přesností na 0,5 kg, zatímco tělesná výška je udávána s přesností na 1 mm. [19]

Norma ČSN EN ISO 15535 obsahuje všeobecné požadavky na tvorbu antropometrických databází. Vzhledem k odlišnosti mezinárodních antropometrických databází, je tato norma jakýmsi stavebním kamenem při jejich tvorbě. Bylo třeba do databáze uvést číslo probanda, jaká ruka byla měřena (levá, pravá), místo bydliště, datum měření, pohlaví (s hodnotami M/F, z anglického a male, a female), datum narození, věk a desetinný věk. Údaje o místě narození, dosaženém vzdělání, věkové skupině a dominantní ruce (s hodnotami L/R, z anglického left, right) byly do databáze také uvedeny. Místo je třeba zapsat ve formátu kódu země podle ISO 3166-1 a místa (např. CZ/Uherský Brod). Datum ve formátu rrrr-mm-dd podle ISO 8601. Ruční výpočet desetinného věku je uveden v příloze D této normy. Pro datum narození i šetření existují tabelované hodnoty, které se přičítají k roku narození a šetření. Hodnoty se od sebe odečtou a zjistí se desetinný věk, který je zaokrouhlen dolů na dvě desetinná místa. Tento věk je důležitý pro rozdělení věku do věkových skupin, který je uveden v příloze E této normy. Dále bylo třeba určit počet osob potřebných pro tvorbu databáze podle metody za-

jišťující, že databáze bude s 95% spolehlivostí obsahovat 5. a 95. percentil populace (v příloze A výše zmíněné normy).

Důležité je, aby databáze splňovala danou formu. Musí se použít kód *ASCII* a údaje musí být od sebe odděleny (například sloupci). Název je vždy uveden v prvním řádku databáze, v anglickém jazyce a dále pak případně v konkrétním jazyce. Všechny rozměry týkající se délek jsou uvedeny v mm, rozměry týkající se váhy v kg. Chybějící údaje musí být označeny symbolem "9999". Norma dále obsahuje kroky pro zjišťování a odstranění chyb měření zahrnující i statistické metody a údaje, který nesmí být opomenuty a také požadavky na antropometrické záznamové listy. Norma je vytvořena tak, aby pomohla tvůrcům při skladbě antropometrické databáze, zamezila co největším možným chybám a byla použitelná pro srovnání s jinými antropometrickými databázemi. [19]

5.2 Předzpracování a čištění dat

Nejprve je nutné si upravit daný datový soubor do formátu, který je strojově čitelný. V našem případě se jedná o soubor aplikace Excel. Poté se práce zaměřuje na nezvyklé či nevyplněné hodnoty.

5.2.1 Příprava datového souboru

Vzhledem k tomu, že aplikace je napsána pomocí jazyka R, je vhodné datový soubor upravit doporučeným způsobem, který lze nalézt v uživatelské příručce viz B.3.1. Tento postup však není podmínkou pro správné zpracování dat.

5.2.2 Čištění dat

Před výpočty a testy je dobré se podívat na jednotlivé proměnné, které daný dataset obsahuje, a tak se zaměřit na tzv. chybné hodnoty. Jako chybnou hodnotu lze považovat odlehlé pozorování (tj. hodnotu, která se nejvíce liší od ostatních), špatně zapsanou hodnotu, duplikovanou hodnotu, chybějící hodnotu apod. Pokud je v datech spatřena chyba, měla by být ručně odstraněna.

V našem případě dataset obsahuje kategorické i numerické proměnné. Vzhledem k tomu, že tyto proměnné vyžadují jiné postupy zpracování, jsou rozlišovány. Numerické proměnné jsou určeny pro statistické metody a výpočty. Zatímco kategorické proměnné určují pouze kategorie, do kterých mohou numerické hodnoty spadat. Vzhledem k tomu, že při tvorbě databáze

je nutné dodržovat určitý postup, např. chybějící údaje označit symbolem 9999 či "NA", nemusí být ošetřeny některé výjimky.

Nicméně malé chyby při vytváření datasetu mohou být zaneseny. Jedním z příkladů může být počet mezer mezi slovy. V našem případě u kategorie pohlaví, byla jedenkrát jedna mezera navíc u znaku skupiny mužů ("M ") a tudíž se vytvořila skupina navíc. Tento nepatrný překlep může mít za následek špatnou kategorizaci, a proto byly všechny mezery odstraněny před a za hodnotami. V aplikaci byla pro to použita funkce *gsub()* pro nahrazení textu.

Dalším problémem jsou odlehle hodnoty. V případě antropometrie rukou se může jednat o náměry u osob, které měly zkrácené šlachy či nebyly schopni natáhnout prsty. Hodnoty lišící se od ostatních mohou zaneść do výpočtů chybu, proto je vhodné tyto hodnoty odstranit. V našem případě byla kontrola odlehlých hodnot prováděna pro numerické proměnné a to podle metod uvedených v kapitole 3.2. V aplikaci jsou vykreslovány boxploty a bagploty. Nejprve bylo pro odstranění odlehlých pozorování aplikováno pravidlo 3-sigma. Hodnoty, které byly mimo interval označující vzdálenost tří směrodatných odchylek od průměru, byly smazány. Poté byla implementována jiná metoda založená na absolutní odchylce mediánů (MAD). Tato metoda je robustnější a lépe detekuje odlehlá pozorování, protože interval je ve vzdálenosti tří směrodatných odchylek od mediánu. Odhad směrodatné odchylky je určen pomocí MAD.

Tyto úpravy jsou prováděny při načítání souboru. Pro přehlednost v aplikaci v záložce *Data Loading* je ponechána informace o tom, kolik hodnot nesplňovalo pravidlo MAD a tudíž byly tyto hodnoty vymazány. Dále je tu ponechána informace o tom, kolik záznamů obsahující neplatnou hodnotu bylo vymazáno. Jedná se o numerické hodnoty, které jsou buďto nevyplněny nebo vyplněny číslem 9999 či symbolem "NA".

5.3 Analýza dat české populace

Tato sekce se zabývá analýzou načtených dat českých probandů. Jsou popsány základní charakteristiky dat, vykresleny grafy pro bližší pochopení použitých statistik, provedeny testy týkající se normality dat a vzájemné vazby mezi proměnnými v závislosti na dané kategorii (věková skupina, pohlaví, BMI, laterita). U všech testů je zvolena hladina významnosti $\alpha = 5\%$. Dále je provedena analýza rozptylu ke zjištění významného rozdílu mezi skupinami a pomocí metody hlavních komponent a shlukování budou vykresleny množiny pro první dvě hlavní komponenty.

K antropometrickému měření rukou došlo v roce 2014-2015, kdy bylo naměřeno 1032 osob v různých populačních skupinách. Nejmladšímu subjektu bylo 17 let a nejstaršímu 83 let, nicméně osoby starší 80 let (2 osoby) a osoby mladší 20 let (mladiství, 14 osob) byly vyřazeny při úpravě datového souboru pro naši studii. Dále bylo pomocí metody MAD odstraněno 42 záznamů (u pravidla 3-sigma to bylo 29 záznamů). Nadále bylo zpracováno 974 údajů.

Bylo ponecháno 6 populačních (věkových) skupin rozdělených po 10 letech, aby byla zajištěna co největší velikost populační skupiny. Populační skupiny jsou tedy následující: 20-29 let, 30-39 let, 40-49 let, 50-59 let, 60-69 let a 70-79 let. Probandi byli rozděleni do těchto skupin podle normy ČSN EN ISO 15535 na základě desetinného věku. Lze se tedy zaměřit na jednotlivou věkovou skupinu probandů a zkoumat vztahy mezi nimi. Přehled lze vidět v tab. 5.1.

Tabulka 5.1: Přehled osob rozdělených do populačních skupin

Věková skupina	Celkový počet	Ženy	Muži
20-29 let	355	109	246
30-39 let	154	58	96
40-49 let	186	97	89
50-59 let	163	88	75
60-69 let	67	43	24
70-79 let	49	37	12
CELKEM	974	432	542

Ve studii bylo zastoupeno téměř 8,6 % leváků (z nichž 2,6 % jsou ženy a 6 % muži) a 91,4 % praváků (z nichž převažují opět muži nad ženami v poměru 5:4). V tab. 5.2 lze vidět přehled.

Tabulka 5.2: Počty leváků a praváků

Dominantní ruka	Celkový počet	Ženy	Muži
levá (L)	84	25	59
pravá (R)	890	407	483

Dále bylo povolání rozděleno podle našeho uvážení na dva hlavní typy, a to na manuální a kancelářskou práci. Při sběru dat se jednalo především o osoby ve výrobě nebo konstruktéry, programátory a další administrativní a technickohospodářské pracovní pozice. Zastoupení manuální práce bylo nepatrně větší než u kancelářské práce viz tab. 5.3.

Tabulka 5.3: Počty osob rozdělených podle kategorie BMI

Typ práce	Celkový počet	Ženy	Muži
kancelářská (K)	445	199	246
manuální (M)	529	233	296

Indexy tělesných hmotností probandů byly rovněž zkoumány. Tento index, většinou značený jako BMI (angl. *body mass index*), se používá pro identifikaci podváhy, nadváhy, normální hmotnosti a obezity. Index se vypočítá jako podíl hmotnosti a kvadrátu výšky osoby. Podle WHO [1] se probandy řadí do 6 skupin (podváha, normální váha, nadváha, mírná obezita, střední obezita a těžká obezita). K tomuto účelu byla vytvořena nová proměnná *BMI kategorie* v původních datech. Z celkového počtu probandů má téměř polovina normální váhu, 1,7 % osob trpí podváhou, 35,4 % nadváhou, 12,6 % mírnou obezitou a jen 1,5 % střední obezitou. Ovšem nevyskytl se ani jeden případ těžké nadváhy, jež může vést k závažným zdravotním problémům. Z tab. 5.4 lze vidět, že muži trpí nadváhou a mírnou obezitou téměř 2x častěji než ženy. Zato počet osob se střední obezitou je u obou vyrovnaný.

Tabulka 5.4: Počty osob rozdělených podle kategorie BMI

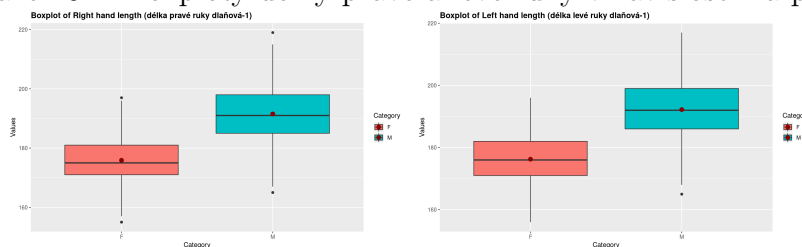
Kategorie	Celkový počet	Ženy	Muži
podváha	17	13	4
ideální váha	474	244	230
nadváha	345	120	225
mírná obezita	123	48	75
střední obezita	15	7	8

5.3.1 Základní charakteristiky antropometrických dat

Tato sekce je věnována výsledkům, které byly získány porovnáním průměrných hodnot. Sledovány byly náměry pravé a levé ruky zvláště v závislosti na pohlaví, dominantní ruce, typu práce, věkové a BMI skupině. V průměru se levé a pravé ruce od sebe příliš nelišily (maximálně o 1 mm). Avšak u obvodu dlaně se rozměry lišily až o 5 mm a u délky palce téměř 2 mm. Stručný přehled všech rozměrů rukou je k dispozici v příloze v tabulkách. B.1 B.2 a B.3. Z důvodu přehlednosti je uveden jen průměr, medián a směrodatná odchylka v závislosti na skupinách. Nicméně aplikace nabízí všechny popisné charakteristiky, které byly uvedeny v kapitole 3.2.1.

Délka ruky ženy je kratší než u muže v průměru o 1,6 cm. Délky i šířky dlaní u žen byly průměrně o 1 cm menší než u mužů. V případě obvodu dlaně se ruce mužů a žen průměrně lišily o 2,5 cm, což bylo předpokládáno, protože muži většinou mají mohutnější ruce než ženy. Ženy i muži mají o 4 mm delší obvod pravé dlaně oproti levé dlani. Délka prostředníku je u žen průměrně kratší o 6 mm a to jak na pravé, tak na levé ruce. Délky ostatních prstů (prostředníku i palce) jsou pro ženy průměrně o 4 až 5 mm kratší. U délky stisku spouště se liší náměry o 7 mm, u palce zhruba o 4 mm. Lze tedy usoudit, že ženy oproti mužům mají celkově menší ruce. Tuto domněnku potvrzují i vykreslené boxploty pro délku ruky v závislosti na pohlaví pro levou i pravou ruku zvlášť viz obr. 5.2.

Obrázek 5.2: Boxploty délky pravé a levé ruky v závislosti na pohlaví



Leváci mají delší ruce než praváci o zhruba 3,5 mm na pravé a 2,5 mm na levé ruce. I u ostatních rozměrů mají leváci větší rozměry než praváci. Leváci mají větší obvod dlaně u své dominantní ruky oproti pravákům o zhruba 6,5 mm, zatímco u pravé ruky je to jen 3 mm. Leváci mají svojí dominantní dlaň širší než pravou. Ovšem v případě obvodu je pravá ruka v průměru větší o 0,1 mm než levá. Jedná se tedy o nepatrný rozdíl. Oproti tomu, praváci mají svou dominantní ruku obvodově o 4 mm větší než levou. To znamená, že praváci mají mohutnější svoji dominantní ruku, zatímco leváci mají obě ruce podobné. Leváci mají rovněž průměrně nepatrně větší ruce zhruba o 1 až 2 mm než praváci. Znamená to tedy, že leváci se přizpůsobili světu pro praváky a mají podobně mohutnou levou i pravou ruku. Rovněž ne vždy jsou rozměry u dominantní ruky kratší než u nedominantní.

U délek rozměrů není pozorován rozdíl mezi typem práce (maximálně se liší o 1 mm). Zajímavé je, že v případě manuální práce je šířka dlaně o 2,5 mm a obvod dlaně o 5 mm větší než u kancelářské. Z toho vyplývá, že manuální práce má vliv na obvod ruky. Zvětšení obvodu ruky může být zapříčiněno buď nadváhou nebo posilováním ruky.

U věkových kategorií se největší rozměry vyskytují ve skupině 30-39 let vyjma délky prostředníku (zde je největší rozměr u kategorie 20-29 let). Nejmenší rozměry jsou u nejstarší skupiny 70-79 let, nicméně u palce je to

na levé ruce skupina 50-59 let. Lze tedy vidět, že rozměry rukou se s věkem snižují průměrně o 5 mm. Největší rozdíl je u obvodu dlaně průměrně o 9,6 mm na pravé i levé ruce. Rozdíly jsou zde patrné.

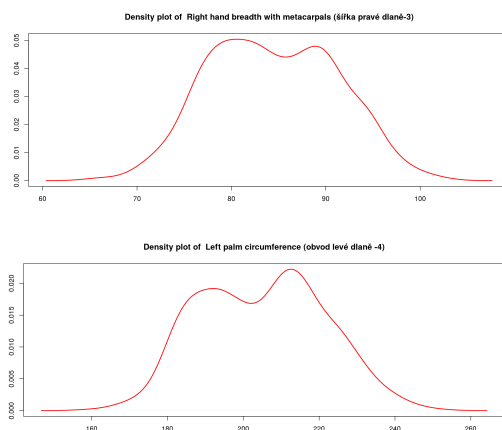
U BMI kategorií jsou vidět rozdíly u obvodu a šířky dlaně mezi osobami trpící podváhou a střední obezitou. Obvod se liší průměrně o 2,8 cm a šířka dlaně se liší na pravé o 9,7 mm na levé o 1,1 cm. I u ostatních rozměrů vyjma délky prostředníku a ukazováku rostou rozměry s narůstajícím BMI (nejmenší rozdíly má podváha, poté ideální váha, nadváha, mírná a nakonec střední obezita).

5.3.2 Testy normality

K zmíněným rozměrům byly provedeny i testy normality. Předpoklad normálního rozdělení je podmínkou pro použití určitých statistických metod (např. dvouvýběrový t-test, jednofaktorová ANOVA, metoda hlavních komponent).

Data byla vizuálně kontrolována pomocí histogramů, odhadů funkcí hustot a Q-Q plotu pro jednotlivé skupiny zvlášť. Odhady funkcí hustoty všech rozměrů (vyjma šířky a obvodu dlaně viz obr. 5.3) se podobá Gaussovo křivce.

Obrázek 5.3: Funkce hustoty pro šířku pravé dlaně a pro obvod levé dlaně



V aplikaci je možno vidět i koeficienty šikmosti a špičatosti, výsledky testů normality a p-hodnoty testů. Výsledky pro jednotlivé rozměry v závislosti na kategoriích jsou k dispozici v příloze B.4, B.5, B.6, B.7, B.8, B.9, B.10 a B.11. S ohledem na počet pozorování je vybrán vhodný test. U většiny skupin je počet pozorování $n > 100$, proto je použit Jarqueho-Berův test. U skupiny 60-69 let, 70-79 let a leváků je $n < 100$ a je zvolen Lilieforsův

test. Pro nízký počet osob trpící podváhou a nadváhou ($n < 20$) je vybrán Sharipův-Wilkův test. Všechny testy testují nulovou hypotézu o tom, zda data pocházejí z normálního rozdělení na hladině významnosti $\alpha = 5\%$. U každého testu jsou uvedeny i p-hodnoty jednotlivých testů. Je-li p-hodnota menší než hladina významnosti α , je zamítnuta nulová hypotéza.

U většiny rozměrů v závislosti na kategoriích nebyla zamítnuta H_0 o tom, že data pocházejí z normálního rozdělení. Nicméně u všech rozměrů, spadající do kategorie kancelářské práce, vyjma délek prstů (palec, prostředník, ukazovák) byly zamítnuty H_0 a to u všech testů. Nicméně koeficienty špičatosti se pohybují okolo hodnoty 2,5, a to jak u pravé tak u levé ruky. Dále u praváků vyjma délky prostředníku a ukazováků byla opět zamítnuta H_0 o normalitě dat pro obě ruce. U délky dlaně poté v kategorii 20-29 let na pravé ruce, ideální váze na levé ruce a nadváze na obou rukou byla H_0 opět zamítnuta. U délky palce u kategorií manuální práce na pravé ruce, žen na obou rukou, praváků na náměrech pravé ruky a nadváze pro levé náměry rovněž byly zamítnuty H_0 u všech testů. U délky pro stisk spouště tomu bylo rovněž tak u typech práce, u praváků a u osob s ideální váhou. Největší počet zamítnutých hypotéz o H_0 bylo u šířky a obvodu dlaně (pohlaví, typ práce, praváci, věková skupina 20-29, 40-49, 50-59 a 70-79 let, osob s ideální váhou, podváhou a mírnou obezitou).

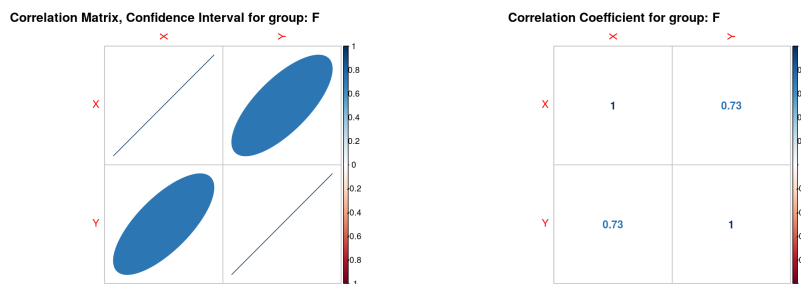
5.3.3 Korelační analýza

Tato část je věnována závislosti délky ruky, šířky dlaně a obvodu dlaně na pohlaví, věku, BMI a typu práce. K tomuto účelu byly využity výpočty korelačních koeficientů (Pearsonův, Kendallův a Spearmanův). Ovšem u některých vazeb může vstupovat ještě další proměnná, která může zkoumané dvě proměnné ovlivňovat. Jedná se tak o klamnou korelaci. Když je závisle proměnná odečtena od zbylých proměnných, je získána čistá korelace. Proto se práce zabývá i parciálními korelačními koeficienty. Přehled lze nalézt v přílohách viz tab. B.13 - B.24.

Koeficienty byly testovány na hladině významnosti $\alpha = 5\%$ a byly porovnávány s p-hodnotou. Je-li p-hodnota menší než hladina významnosti α , je zamítnuta H_0 o nulovosti korelačního koeficientu, a tedy koeficient lze považovat za významný. Pro vizualizaci bylo kromě bodového grafu s vykreslením lineární trendu použito i grafické zobrazení korelační matice a konfidenčních intervalů, kdy tmavší barvy označují závislost mezi proměnnými jako například na obr. 5.4.

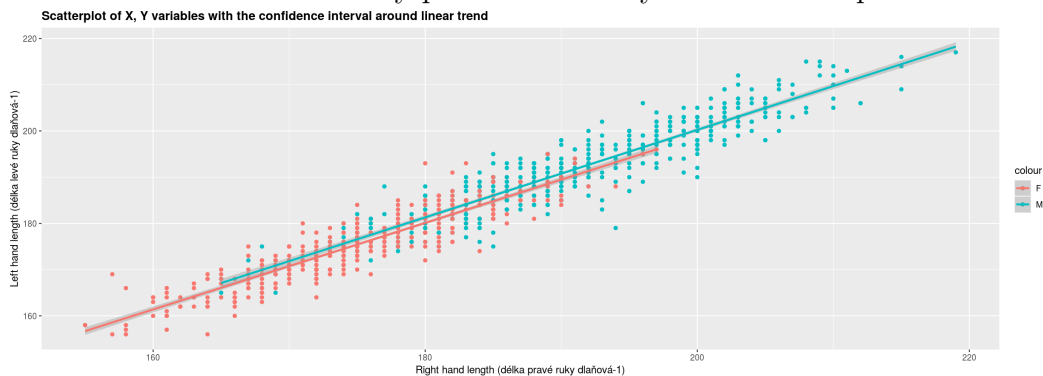
Vizualizace dat na obr. 5.5 naznačuje, že délka pravé a levé ruky na sobě závisí a že ženy mají kratší ruce než muži. Pearsonův korelační koeficient lze

Obrázek 5.4: Korelační matice a konfidenční interval pro pravý palec - ženy



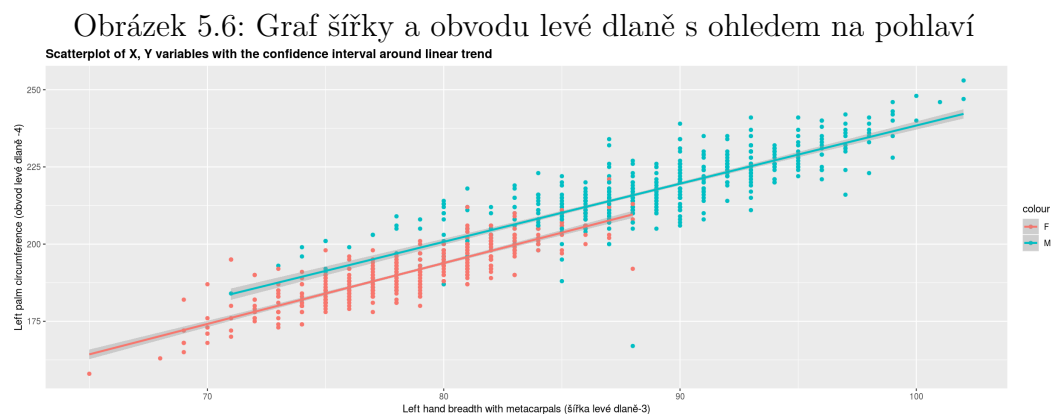
považovat za významný pro ženy i muže a jeho hodnota je 0,92. Rovněž délka pravé a levé dlaně spolu souvisí (koeficient je 0,83 pro ženy a 0,87 pro muže a lze považovat závislost za silnou). Podobně tomu tak bylo i u ostatních rozměrů. Šířka pravé a levé dlaně vykazovala Pearsonův koeficient 0,8, obvod pravé a levé dlaně 0,9. U délek prstů se ovšem skupiny mužů a žen téměř překrývaly. Ale u mužů i žen existovala závislost mezi rozměry délek prstů pravé a levé ruky (Pearsonův korelační koeficient je v rozmezí 0,7 - 0,8). Rozměry pravé a levé ruky tedy na sobě závisí (když je velká jedna ruka, tak i druhá).

Obrázek 5.5: Graf délky pravé a levé ruky s ohledem na pohlaví



Významná korelace byla pozorována u i šířky dlaně a délky dlaně pravé a levé ruky pro ženy a muže zvlášť. Pearsonův korelační koeficient se pohyboval okolo 0,8 a je statisticky významný. Obdobně tomu bylo u šířky dlaně a délky ruky, nicméně korelační koeficienty se pohybovaly okolo 0,5 u mužů i žen. Dále byl zkoumán vztah, jestli šířka dlaně souvisí s obvodem. Tato domněnka byla potvrzena, Pearsonův korelační koeficient je 0,5 u žen a 0,47 u mužů na levé a 0,48 u žen a 0,5 u mužů na pravé ruce. Existuje tedy korelace mezi šířkou a délkou dlaně, menší závislost mezi šířkou dlaně

a délkou ruky a mezi délkou ruky a obvodem dlaně. Korelační koeficienty lze považovat za statisticky významné. Ženy měly opět menší rozměry než muži viz obr. 5.6. Dalším zkoumaným vlivem byl typ práce. V případě obvodu a šířky dlaně byl Pearsonův korelační koeficient vysoký (0,9) lze ho považovat za statisticky významný. U ostatních porovnávaných rozměrů byl nižší.



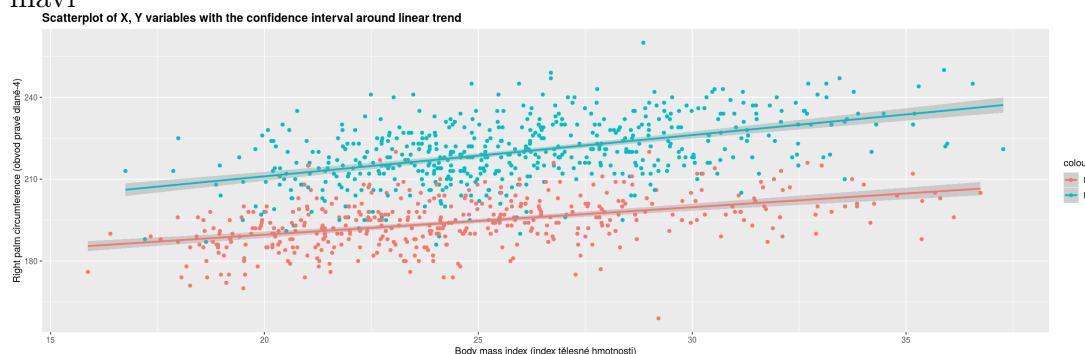
Dalším předmětem zkoumání bylo, zda s věkem roste délka ruky. Jsou uvedeny Pearsonovy korelační koeficienty. S věkem jsou šlachy, vazy opotřebovanější, tudíž by se rozměry mohly zmenšovat. Byla zpozorována významnější závislost mezi věkem a délkou palce (0,19), mezi věkem a šířkou dlaně (0,31 u žen, 0,25 u mužů na pravé ruce, 0,2 na levé ruce) a mezi věkem a obvodem dlaně (0,2). Lze tedy říci, že s věkem nepatrně zvětšuje šířka a obvod dlaně, délka palce a že ženy mají menší rozměry než muži.

Dále bylo potvrzeno, že čím jsou lidé vyšší, tím mají větší délky rukou (Pearsonův korelační koeficient se pohyboval okolo 0,6), dlaní (0,5), obvod (0,28 u žen a 0,34 u mužů), délku prstů (ukazovák 0,5 u žen a 0,46 u mužů, palec 0,37 u žen a 0,28 u mužů, prostředníku 0,47 u žen a 0,53 u mužů). Pearsonovy korelační koeficienty lze považovat za významné.

S rostoucím BMI by měla růst i šířka a obvod dlaně vzhledem k tomu, že lidé trpící obezitou jsou celkově mohutnější. Pearsonovy korelační koeficienty šířky dlaně se u žen pohybovaly okolo 0,33 a u mužů 0,43. Hodnota Pearsonova koeficientu u obvodu dlaně také nepatrně rostla pro ženy (0,45) i muže (0,48), avšak index BMI žen byl menší než u mužů viz obr. 5.7. Byl zkoumán i vliv typ práce na zmíněné rozměry. Pearsonovy korelační koeficienty se pohybovaly průměrně okolo 0,4. Zamítnutí nulové hypotézy svědčí o významnosti koeficientů.

Vzhledem k tomu, že rozměry pravé i levé ruky rostly s výškou osoby, může být výška negativním faktorem při porovnání (korelační koeficienty

Obrázek 5.7: Graf závislosti BMI na obvodu pravé dlaně s ohledem na pohlaví



byly vysoké). Proto byla tato informace zahrnuta k porovnání levých a pravých náměrů (viz výše). I přes eliminaci klamné korelace zůstaly parciální korelační koeficienty u všech náměrů signifikantní (Pearsonovy parciální korelační koeficienty se pohybují v rozmezí 0,68 - 0,89).

5.3.4 Analýza rozptylu (ANOVA)

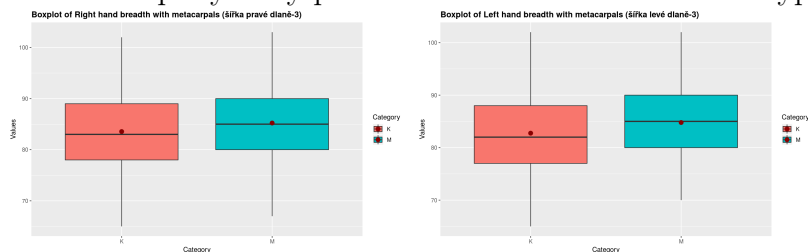
Tato část se zabývá statisticky významnými rozdíly mezi pohlavím, věkovými skupinami, povoláním, dominancí ruky a BMI kategorií pro rozměry rukou. Vzhledem k tomu, že skupiny mají kategorie různých délek (pohlaví, dominance ruky a typ práce dvě, věková skupina šest, BMI kategorie pět kategorií) je třeba použít i správnou metodiku. Jak již bylo popsáno v kapitole 3.3.4, je třeba kontrolovat i předpoklady (normalitu, test pro shodu rozptylů) a použít buďto parametrický nebo neparametrický test. Pro testování normality byl použit Jarque-Bera test, nicméně lze vycházet z výsledků uvedených v 5.3.2. Rozhodování o shodě dvou průměrů (resp. mediánů) bylo na základě velikosti p-hodnot příslušných testů. Mnohonásobné porovnávání bylo rovněž provedeno na základě párového srovnání Tukeyho testu a pořadového Wilcoxonova testu, kde byla použita Holmovo korekce p-hodnoty, aby se při mnohonásobném porovnání nezvětšovala chyba I. druhu. U tohoto testu se rovněž vypisují v aplikaci p-hodnoty. Přehled výsledků lze nalézt v příloze viz tab. B.25, B.26, B.27, B.28 a B.29.

Muži mají delší ruce než ženy, jak lze vidět na obr. 5.2, měl by tedy existovat významný rozdíl mezi muži a ženami. Vzhledem k tomu, že v případě délky ruky odpovídají data mužů i žen normálnímu rozdělení, ale rozptyly se liší, byl použit Welchův t-test o shodě průměrů pro různé rozptyly. P-hodnota testu pravé i levé ruky vyšla menší než 0,001, a proto byla zamít-

nuta nulová hypotéza o shodě průměrů a tím lze považovat rozdíl mezi muži a ženami v případě délky prstů na pravé i levé ruce statisticky významný. U šířky dlaně test normality vyšel jen pro ženy, a shodnost rozptylů byla zamítnuta, proto bylo lepší použít Wilcoxonův pořadový test, jež opět nezamítl hypotézu o rozdílnosti dvou skupin. Obdobně tomu bylo i u obvodu dlaně. Testy byly provedeny pro všechny rozměry a lze rozdíly u pohlaví považovat za statisticky významné.

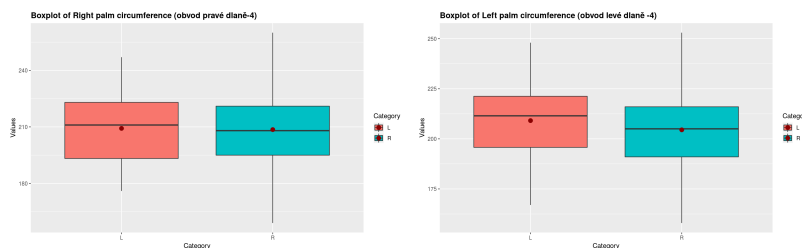
U kancelářské a manuální práce nebyl zaznamenán až tak velký rozdíl u jednotlivých rozměrů, rozdíly nastaly u šířky a obvodu dlaně, délky palce. U délky ruky nesplňovaly data kancelářské práce normální rozdělení, rovněž nebyl splněn předpoklad shody rozptylů, a tak byl použit pořadový Wilcoxonův test. Nebyl prokázán rozdíl mezi typem práce. U šířky a obvodu dlaně byl opět porušen předpoklad normality dat a pořadový Wilcoxonův test nezamítl hypotézu o rozdílu mezi skupinami pro pravou i levou ruku. Další rozdíl byl u délky palce. U ostatních rozměrů nebyl prokázán rozdíl mezi typem práce.

Obrázek 5.8: Boxploty šířky pravé i levé dlaně v závislosti na typu práce



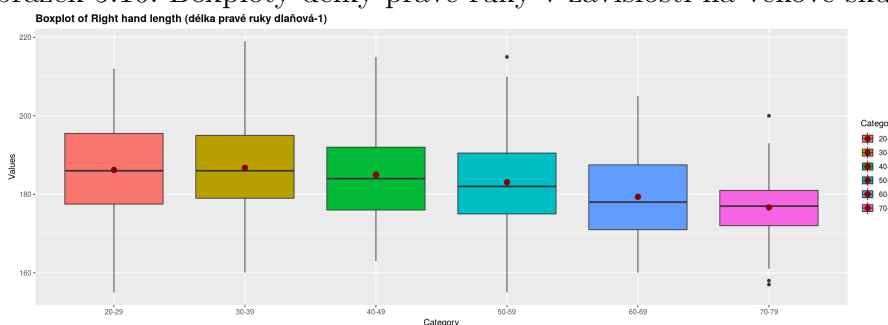
Dále byl zpozorován rozdíl mezi praváky a leváky. U praváků nebyl splněn předpoklad normality u všech rozměrů vyjma délky ukazováku a prostředníku. Pořadový Wilcoxonův test zaznamenal rozdíl mezi dominantní rukou u pravé ruky, ale u levých náměrů ne. U levé ruky byl zaznamenán rozdíl jen u délky ruky, obvodu a šířce dlaně, u pravé ruky u délky pro stisk spouště a délky prostředníku). Z toho vyplývá, že praváci mají šířku a obvod dlaně u své dominantní ruky podobný jako leváci na své nedominantní ruce. To stejné platí i pro délku ruky. Zatímco u leváků nebyly zaznamenány rozdíly mezi svou dominantní rukou a nedominantní rukou, u praváků byly odlišnosti. Lze z toho usoudit, že leváci svou nedominantní ruku používají častěji než praváci.

Obrázek 5.9: Boxploty obvodu pravé i levé dlaně v závislosti na dominanci ruky



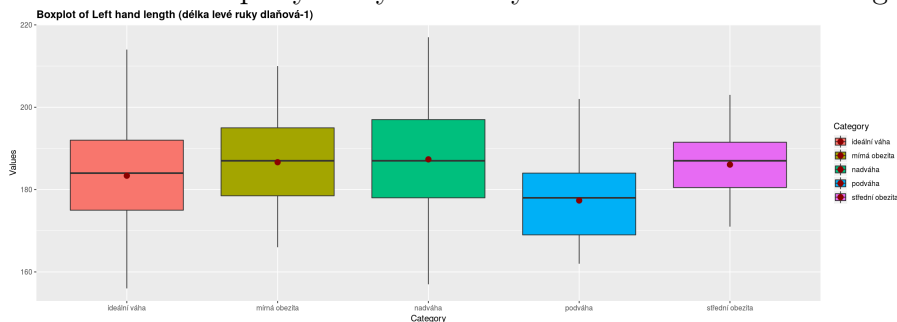
U věkových skupin nebyl splněn předpoklad normality u délky pravé dlaně u skupiny 20-29 let, u šířky dlaně a délky prostředníku navíc skupina 40-49 let, u obvodu navíc skupina 50-59 let, 70-79 let, u délky pro stisk spouště pravé ruky jen u kategorie 70-79 let. Byly tedy použity přístupy jednofaktorové analýzy rozptylu. U délky pravé ruky byly potvrzeny rozdíly mezi skupinami (lze vidět i na obr. 5.10). Byl použit Tukey test pro porovnání skupin mezi sebou a bylo zjištěno, že největší rozdíl byl mezi nejstarší skupinou 70-79 let a 30-39 let. Druhý největší rozdíl byl mezi nejstarší a nejmladší skupinou 20-29 let. V případě obvodu a šířky dlaně byl proveden Kruskalův-Wallisův test a nebyly zamítnuty statisticky významné rozdíly mezi skupinami. Významné rozdíly byly i u ostatních rozměrů vyjma délky levého palce. Většinou největší rozdíly byly vidět u nejstarší skupiny 70-79 let a 30-39 let.

Obrázek 5.10: Boxploty délky pravé ruky v závislosti na věkové skupině



Poslední zkoumaná skupina byla BMI skupina. U všech rozměrů vyjma délky levého prostředníku byl zaznamenán rozdíl mezi BMI kategoriemi. Rozdíly byly patrné mezi podváhou a ostatními kategoriemi, největší mezi podváhou a nadváhou na pravé i levé ruce. To naznačuje i vizualizace dat viz obr. 5.11.

Obrázek 5.11: Boxploty délky levé ruky v závislosti na BMI kategorii



5.3.5 Metoda hlavních komponent a shlukování

Tato část zkoumá, zda lze pomocí shluků identifikovat povolání nebo BMI kategorii. K dispozici je 8 rozměrů ruky pro pravou i levou ruku zvlášť. Z vybraných rozměrů je možno pomocí korelační matice zjistit, které rozměry se nejvíce ovlivňují. Cílem je, aby pomocí metody hlavních komponent (PCA) byla zredukována dimenze úlohy a poté byly vytvořeny shluky a byla přidána informace o kategorii. Lze tak zjistit, zda pomocí rozměrů se dá identifikovat typ povolání nebo BMI kategorie.

Na začátku rovněž musí být splněn předpoklad použití metody PCA. Data musí pocházet z normálního rozdělení, a proto je na začátku kontrolován předpoklad nulové střední hodnoty a jednotkového rozptylu. Pokud předpoklad není splněn, provede se škálování. Pomocí vizualizace Paretova grafu lze mít představu o tom, kolik hlavních komponent vysvětlí přinejmenším 80 % rozptylu.

Vzhledem k tomu, že je tato část zpracována grafickou metodou, jsou uvažovány dvě hlavní komponenty. Pro dvě hlavní komponenty se vytvoří shluky a obarví body podle povolání a BMI kategorie. Je tak k dispozici informace o tom, jak by vypadaly teoretické shluky a jak vypadají shluky vytvořené z původních dat.

Nejprve je zkoumána korelační matice. V tab. 5.5 lze zpozorovat, že existují silně korelované proměnné. Délka ruky a délka dlaně na sobě silně závisí stejně jako šířka a obvod dlaně. Všechny rozměry rukou jsou ponechány, aby nebyla ztracena informace eliminací proměnných.

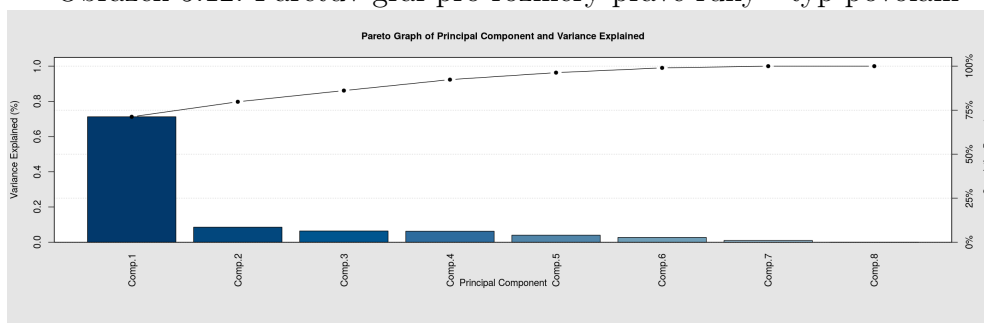
Pomocí metody PCA jsou zjištěny hlavní komponenty. První komponenta vysvětlí 71,3 % a druhá komponenta 8,5 % celkového rozptylu. Stačí tedy dvě hlavní komponenty k tomu, aby vysvětlily 79,8 % variability dat. Grafická vizualizace Paretova grafu viz obr. 5.12 naznačuje, že tyto dvě komponenty stačí k popsání rozměrů celé ruky.

Jsou k dispozici tedy dvě hlavní komponenty, jež vznikají ortogonální

Tabulka 5.5: Korelační matice - pravá ruka

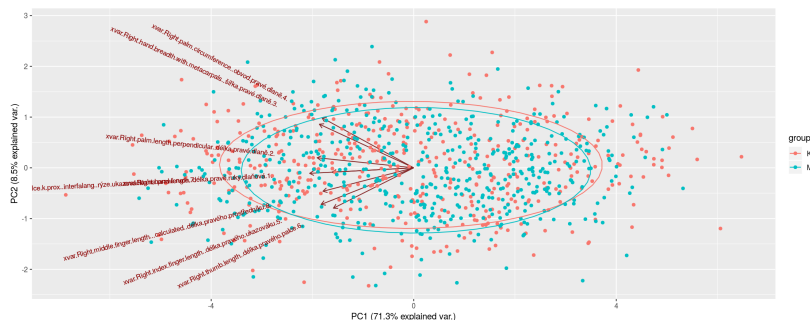
Pravé náměry	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Délka ruky (1)	1	0,92	0,76	0,75	0,8	0,63	0,67	0,88
Délka dlaně (2)	0,92	1	0,74	0,72	0,69	0,59	0,65	0,63
Šířka dlaně (3)	0,76	0,74	1	0,92	0,63	0,57	0,59	0,62
Obvod dlaně (4)	0,75	0,72	0,92	1	0,59	0,51	0,55	0,62
Délka ukazováku (5)	0,8	0,69	0,63	0,59	1	0,68	0,59	0,76
Délka palce (6)	0,63	0,59	0,57	0,51	0,68	1	0,51	0,55
Délka pro stisk spouště (7)	0,67	0,65	0,59	0,55	0,59	0,51	1	0,55
Délka prostředníku (8)	0,88	0,63	0,62	0,62	0,76	0,55	0,55	1

Obrázek 5.12: Paretův graf pro rozměry pravé ruky - typ povolání



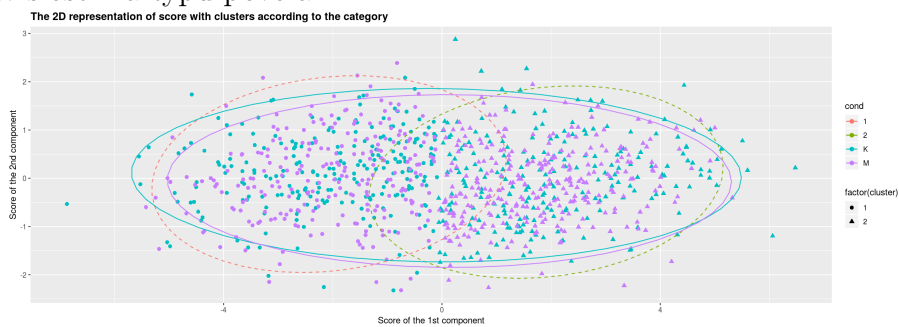
lineární kombinací původních dat do 2-rozměrného prostoru. V grafu 5.13 je zobrazen bodový graf první a druhé komponenty s konfidenčními množinami tvořených podle skupin a směry (původními rozměry rukou). Vzhledem k tomu, že data odpovídají normálnímu rozdělení, konfidenční množiny mají tvar elipsy. Data by měla ležet s 95% pravděpodobností uvnitř elipsy. Také lze vidět, že v našem případě nelze s jistotou určit rozdělení kancelářské a manuální práce, protože skupiny se překrývají.

Obrázek 5.13: Závislost první a druhé hlavní komponenty v závislosti na typu povolání



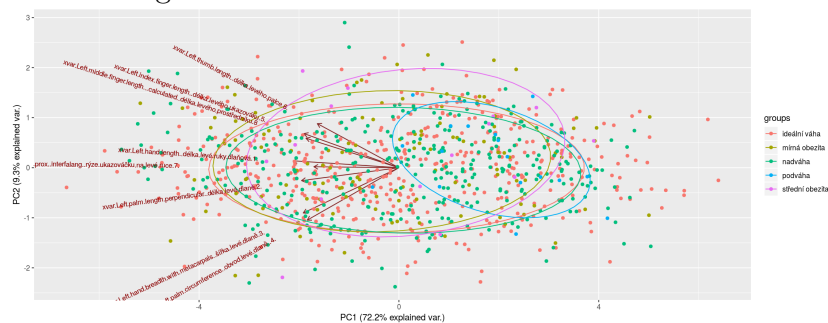
Je-li přidána informace o počtu shluků (v našem případě 2) a použita metoda K means, lze rozdělení dat identifikovat podle shluků a barevně jsou odlišeny body jednotlivých kategorií. Zároveň jsou vykresleny konfidenční elipsy t-rozdělení a normálního rozdělení (čárkovaně). Lze vidět, že elipsy t-rozdělení jsou větší, protože je v nich informace o rozptýlenosti původních dat (normováním je ztracena informace o variabilitě). V tomto případě elipsy neodpovídají barvám kategorií původních dat viz graf 5.14, proto nelze na základě rozměrů identifikovat typ povolání.

Obrázek 5.14: Závislost dat první a druhé hlavní komponenty se shluky v závislosti na typu povolání



Obdobně tomu bylo u BMI kategorií a to v případě pravé i levé ruky. Nicméně z grafu 5.15 lze usoudit, že existují odlehlá pozorování, která neleží v konfidenční množině normálního rozdělení. Nelze identifikovat podle rozměrů BMI kategorie. Jedním z důvodů, že nebyla identifikace pomocí metody PCA průkazná, může být i to, že nebyl splněn předpoklad normality pro všechny vstupní data.

Obrázek 5.15: Závislost první a druhé hlavní komponenty se shluky v závislosti na BMI kategoriích - levá ruka



Rovněž bylo zjištěno, že pomocí délky ruky a šířky dlaně lze vysvětlit 88 % celkové variability dat v první komponentě viz tab.5.6.

Tabulka 5.6: Rozptyl hlavních komponent

	Komp.1	Komp.2
Rozptyl	1,761	0,24
Směrodatná odchylka	1,327	0,489
Poměr celkového rozptylu	0,880	0,119
Kumulující poměr	0,880	1

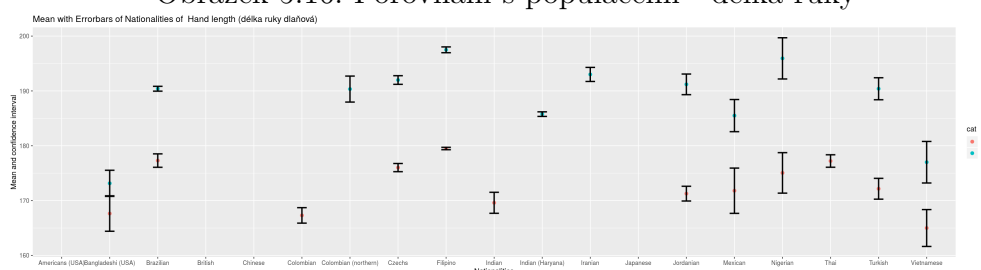
5.4 Porovnání s ostatními populacemi

Důležitou součástí této práce bylo porovnání výsledků české populace s ostatními národy, které se zabývaly tématem antropometrie rukou. K tomuto účelu byla vytvořena databáze z dostupných výsledků ostatních populací. Konkrétně se jedná o obyvatelky Ameriky [29], Anglie [29], Číny [29], Japonska [29] a Thajska [11] a dále pak o obyvatele Bangladéše žijící v Americe [25], Filipín [15], Indie ([32], [12]), Íránu [31], Jordánska [29], Mexika [29], Turecka [11], Brazílie [14], Nigérie ([34], [6]), Kolumbie [35] a Američany vietnamského původu [24]. Údaje o rozměrech ruky pro tyto národnosti byly publikovány ve vědeckých článcích, odkud byly získány informace o průměru (*Mean*), směrodatné odchylce (*SD*) a počtu pozorování (*n*). Vzhledem k tomu, že k dispozici nebyly všechny rozměry, které byly naměřeny v naší studii, byla porovnávána jen dlaňová délka ruky (1), délka dlaně (vypočtená z dlaňové délky ruky a délky prostředníku, 2), šířka dlaně ve výši metacarpalia (3) a délka prostředníku (8). U brazilské studie chyběla informace o délce prostředníku, a tak byla vypočtena z délky ruky a délky dlaně. Některé studie (Indové ze státu Harijána, Iránci a Kolumbijci) obsahovaly všechny zmíněné rozměry. Přehled je k dispozici v příloze B.30, tab. B.31. Pro českou populaci byly hodnoty rozměrů udány jako průměrné hodnoty pravých a levých náměrů zkoumaných rozměrů.

Jsou k dispozici jen omezené informace, rovněž u rozměrů byl doložen průměr (není tedy k dispozici informace o směrodatné odchylce), je provedeno porovnání vizuálně pomocí 95% oboustranných intervalů spolehlivosti při známém rozptylu pro každou populaci pomocí vzorce $(\bar{x} - t_{1-\alpha/2} \frac{s(x)}{\sqrt{n}}; \bar{x} + t_{1-\alpha/2} \frac{s(x)}{\sqrt{n}})$, kde $t_{1-\alpha/2}$ je kvantil Studentova rozdělení s $n - 1$ stupni volnosti a hladinou $1 - \alpha/2$. Ponechána je i informace o t-statistice Welchova t-testu pro nestejný rozptyl a rozdíl zvolené národnosti (v našem případě české populace) s ostatními populacemi pro ženy a muže zvlášť. Národy, které nevedly informaci o průměru a rozdílu, které nepřesáhly výrazného rozdílu nejsou diskutovány.

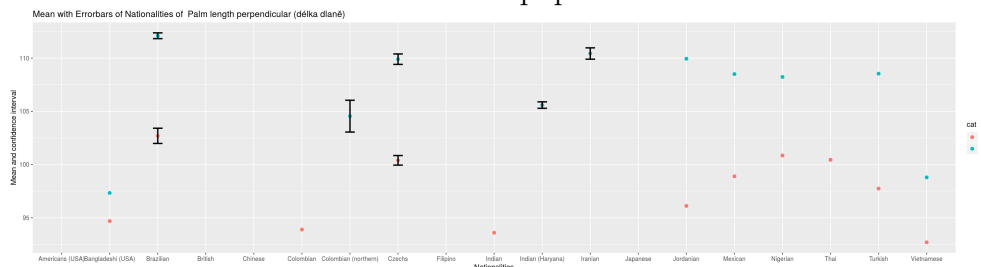
Na obr. 5.16 lze spatřit porovnání populací u rozměru délky ruky pro muže a ženy dohromady. Největších rozdílů bylo dosaženo u žen porovnáním s Vietnamkami, jejichž délka ruky je průměrně o 1,1 cm menší než u českých žen. Kolumbijky mají menší ruce o 8,72 mm, obyvatelky Bangladéše o 8,38 mm, Indky o 6,42 mm, Jordánky o 4,75 mm, Mexičanky o 4,22 mm, obyvatelky Turecka o 3,86 mm. Filipínky mají větší ruce o 3,48 mm, Brazilky o 1,28 mm a Thajky o 1,2 mm. U mužů byl největší rozdíl oproti Bangladéšanům, jež mají o 1,8 cm menší ruce než Češi. Vietnamci mají o 1,5 cm menší ruce, Mexičani o 6,5 mm a Indové o 6,22 mm. Filipínci mají větší ruce o 5,5 mm, Nigerijky o 4 mm. Lze tedy říct, že obyvatelé Bangladéše a Vietnamu mají výrazně menší ruce než Češi.

Obrázek 5.16: Porovnání s populacemi - délka ruky



U délky dlaně bylo třeba hodně hodnot dopočíst, proto nejsou uvedeny směrodatné odchylky. Na obr. 5.17 lze vidět intervalové porovnání jen se čtyřmi národy, u zbylých je uvedena informace o průměru. U žen největšího rozdílu bylo s Vietnamkami, jež mají menší dlaně o 7,7 mm, dále pak Indky o 6,8 mm, Kolumbijky o 6,5 mm, obyvatelky Bangladéše o 5,7 mm a Jordánky o 4,3 mm. Muži dosahovali největších rozdílů opět s Bangladéšany o 1,3 cm, s Vietnamci o 1,1 cm, s Kolumbijci o 5,4 mm a Indy o 4,3 mm, jež mají menší délku dlaně.

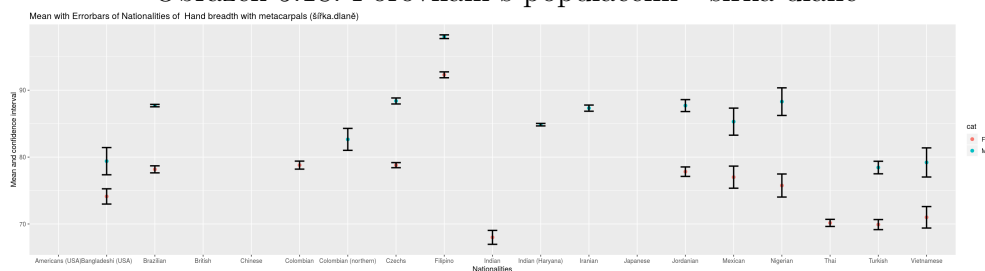
Obrázek 5.17: Porovnání s populacemi - délka dlaně



U šířky dlaně bylo největších rozměrů dosaženo u Filipínců, kdy Filipínky mají o 1,3 cm a Filipínci o 9,6 mm větší než Češi. Ve všem ostatních

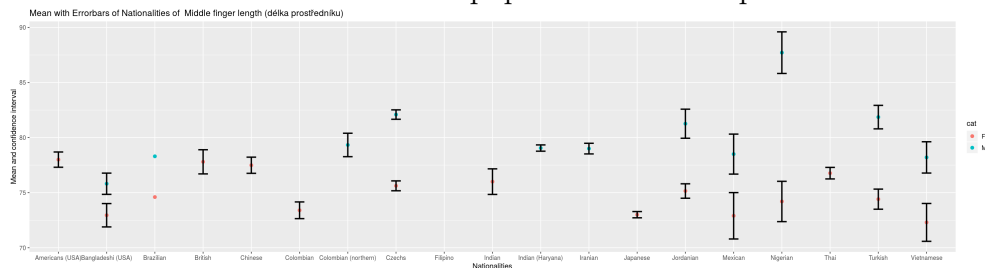
případech mají Češky větší šířku dlaně než ostatní národnosti např. oproti Indkám o 10,8 mm, Turkyním o 8,9 mm, Thajkám o 8,64 a Vietnamkám o 7,8 mm. U mužů jsou rozměry oproti Turkům o 1 cm, Vietnamcům o 9,2 mm, Bangladéšanům o 9 mm a Kolumbijcům o 5 mm menší než u Čechů, jak můžete vidět na obr. 5.18.

Obrázek 5.18: Porovnání s populacemi - šířka dlaně



Posledním porovnávaným rozměrem je délka prostředníku. Z obr. 5.19 lze usoudit, že největší rozdíly u žen jsou s porovnáním s Vietnamkami o 3,32 mm, které mají menší prostředník a s Američankami o 2,4 mm prostředník větší. Dále pak Nigerijci mají o 5,62 mm větší prostředník než Češi, zatímco Bangladéšané mají prostředník o 6,3 mm menší prostředník.

Obrázek 5.19: Porovnání s populacemi - délka prostředníku



Je-li shrnuto výše uvedené, největších rozdílů je dosaženo porovnáním s obyvateli Bangladéše, Filipín, Vietnamu. Bangladéšané a Vietnamci mají menší rozměry než Češi a Filipínci větší rozměry než Češi. Další významné rozdíly jsou pozorovány u Indů a Turků. Bohužel pro naše porovnání nebyly nalezeny informace o obyvatelích Evropské Unie vyjma Německa, pro něhož bylo porovnání již provedeno ve studii viz [10].

6 Závěr

Cílem této práce byla tvorba statistického nástroje, který bude nápomocný řadě výzkumných pracovníků a zájemců o statistické zpracování dat. Díky němu je možné snadno provést exploratorní analýzu, která je ideálním vstupním bodem pro bádání v pozorovaných datech (např. v oblasti *Data Science*).

K tomuto účelu byl vybrán open-source programovací jazyk R, jež je pro statistická zpracování vhodný. Za pomoci knihovny *Shiny* byla vytvořena aplikace v prostředí RStudio. Díky platformě *shinyapps.io* byla aplikace vystavena na webu (<https://datatesting.shinyapps.io/final/>). Do aplikace je možné nahrát soubor ve formátu *.xlsx* a *.xls*. Ostatní formáty nejsou podporovány. Při načítání souboru se provádí kontrola na odlehlá pozorování a neplatné hodnoty. Záznamy s chybějícími údaji jsou vymazány. Rovněž tak odlehlá pozorování, které jsou detekovány pomocí absolutní mediánové odchylky (MAD). Výsledky statistických metod aplikované na uživatelem vybrané proměnné jsou logicky uspořádány do strukturovaných kategorií a sekcí. K dispozici jsou základní popisné charakteristiky, tabulky četností, přehledové tabulky, histogramy, boxploty, bodové grafy, testy normality, zjištění závislosti mezi proměnnými pomocí korelačních koeficientů (resp. partiálních korelačních koeficientů). Dále je k dispozici analýza rozptylu pro zjištění, zda existuje statisticky významný rozdíl mezi skupinami, metoda hlavních komponent a shlukování.

Pomocí aplikace byla prozkoumána data antropometrie rukou. Ve studii bylo naměřeno celkem 1032 osob, zpracováno bylo 974 hodnot. Zastoupení 8,6 % leváků a 91,4 % praváků odpovídá reálnému procentuálnímu zastoupení v populaci České republiky. Byly zkoumány statisticky významné rozdíly mezi skupinami. Rozdíl mezi muži a ženami byl zřejmý. Největší rozdíl lze přirozeně vidět u obvodu dlaně, kde se rozměry lišily o 2,5 cm. Rozdíl mezi praváky a leváky nebyl statisticky významný u většiny měřených rozměrů. Obvod dlaně leváků vykazuje minimální rozdíl mezi dominantní anedominantní rukou. Zatímco u praváků byl pozorován rozdíl 4 mm mezi obvodem pravé a levé ruky. To může poukazovat na to, že leváci používají obě ruce stejně často a to pravděpodobně z důvodu, že většina nástrojů je určena pro praváky. Rozdíly mezi kancelářskou a manuální prací nebyly výrazně odlišné. U věkových skupin se rozměry většinou zmenšují s narůstajícím věkem. Rozdíly mezi jednotlivými BMI kategoriemi byly považovány za statisticky významné. Byla pozorována i závislost mezi výškou a délkou ruky a dále pak mezi rozměry rukou. Korelační koeficienty zůstaly statisticky

významné i po odstranění klamné korelace. Krátká kapitola byla věnována rozeznání skupiny na základě rozměrů rukou pomocí metody hlavních komponent a shlukování. Na základě výsledků však nebylo možné identifikovat zadané skupiny.

Poslední část práce se týkala porovnání výsledků s ostatními populacemi. Za tímto účelem byla v aplikaci vytvořena funkce vyžadující vstupní data s konkrétně danou strukturou. Porovnáním české populace s ostatními bylo zjištěno, že největšího rozdílu je dosaženo s obyvateli Bangladéše, Filipín a Vietnamu.

Literatura

- [1] *Body mass index - BMI* [online]. [cit. 2019/05/05]. Dostupné z: <http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi>.
- [2] *Evropské normy* [online]. Vaše Evropa: Evropská unie, 2018. [cit. 2019/02/10]. Dostupné z: https://europa.eu/youreurope/business/product/standardisation-in-europe/index_cs.html.
- [3] ISO 7250: Basic human body measurements for technological design, 1998.
- [4] *Best Practices in Preparing Data Files for Importing into R* [online]. STHDA: Statistical tools for high-throughput data analysis. [cit. 2019/03/29]. Dostupné z: <http://www.sthda.com/english/wiki/best-practices-in-preparing-data-files-for-importing-into-r>.
- [5] ABDI, H. – WILLIAMS, L. Tukey's Honestly Significant Difference (HSD) Test. 05 2010.
- [6] ADELEYE, A. A. – AKANBI, O. G. Hand cumulative trauma disorders in Nigerian custom tailors: the need for redesign of manual scissors. *Ergonomics*. 2015, 58, 8, s. 1410–1423. doi: 10.1080/00140139.2015.1012123.
- [7] ANDEL, J. *Statistické metody*. Matfyzpress, Praha, 2003. ISBN 80-85863-27-8.
- [8] BALAKRISHNAN, V. – YEOW, P. H. Hand anthropometry and SMS satisfaction. *Journal of Applied Sciences*. 05 2008, 8, 5, s. 816–822. doi: 10.3923/jas.2008.816.822.
- [9] BAYRAKTAR, N. K. – ÖZŞAHİN, E. Anthropometric measurement of the hand. *Eastern Journal of Medicine*. 2018, 23, 4, s. 298. doi: 10.5505/ejm.2018.03164.
- [10] BURES, M. – GORNER, T. – SEDIVA, B. Hand anthropometry of Czech population. In *Industrial Engineering and Engineering Management (IEEM), 2015 IEEE International Conference on*, s. 1077–1082. IEEE, 12 2015. doi: 10.1109/IEEM.2015.7385814.
- [11] ÇAKIT, E. et al. A survey of hand anthropometry and biomechanical measurements of dentistry students in Turkey. *Human Factors and Ergonomics in Manufacturing & Service Industries*. 11 2014, 24, 6, s. 739–753. doi: 10.1002/hfm.20401.

- [12] CHANDRA, A. – CHANDNA, P. – DESWAL, S. Analysis of hand anthropometric dimensions of male industrial workers of Haryana State. *International Journal of Engineering*. 01 2011, 5, 3, s. 242–256.
- [13] CYHELSKÝ, L. – KAHOUNOVÁ, J. – HINDLS, R. *Elementární statistická analýza*. Management Press, 1999. ISBN 80-7261-003-1.
- [14] SILVA, G. V. – HALPERN, M. – GORDON, C. C. Anthropometry of Brazilian air force pilots. *Ergonomics*. 2017, 60, 10, s. 1445–1457. doi: 10.1080/00140139.2017.1301575.
- [15] DEL PRADO-LU, J. L. Anthropometric measurement of Filipino manufacturing workers. *International Journal of Industrial Ergonomics*. 06 2007, 37, 6, s. 497–503. doi: 10.1016/j.ergon.2007.02.004.
- [16] DOMO, I. *Data Never Sleeps 6.0* [online]. 2018. [cit. 2019/03/29]. Dostupné z: <https://www.domo.com/learn/data-never-sleeps-6>.
- [17] ŠEDIVÁ, B. *Mnohorozměrné statistické metody: ANOVA* [online]. 2013/2014. [cit. 2019/05/05]. Dostupné z: <https://portal.zcu.cz/portal/studium/courseware/kma/msm/prednasky.html>.
- [18] FALLAHI, A. – JADIDIAN, A. The effect of hand dimensions, hand shape and some anthropometric characteristics on handgrip strength in male grip athletes and non-athletes. *Journal of human kinetics*. 11 2011, 29, s. 151–159. doi: 10.2478/v10078-011-0049-2.
- [19] STANDARDIZATION, I. O. ISO 15535: General Requirements for Establishing Anthropometric Databases, 1996.
- [20] GARCÍA-CÁCERES, R. G. et al. Hand anthropometry of the Colombian floriculture workers of the Bogota plateau. *International Journal of Industrial Ergonomics*. 2012, 42, 2, s. 183–198. doi: 10.1016/j.ergon.2011.12.002.
- [21] GIBILISCO, S. *Statistika bez předchozích znalostí*. Computer press, 2009. ISBN 987-80-251-2465-9.
- [22] HEBÁK, P. et al. *Statistické myšlení a nástroje analýzy dat*. Informatorium, 2013. ISBN 978-80-7333-105-4.
- [23] HENDL, J. *Přehled statistických metod zpracování dat: analýza a metaanalýza dat*. Portál, sro, 2004. ISBN 978-80-262-0981-2.
- [24] IMRHAN, S. N. – NGUYEN, M.-T. – NGUYEN, N.-N. Hand anthropometry of Americans of Vietnamese origin. *International Journal of Industrial Ergonomics*. 12 1993, 12, 4, s. 281–287. doi: 10.1016/0169-8141(93)90098-X.

- [25] IMRHAN, S. N. – SARDER, M. – MANDAHAWI, N. Hand anthropometry in Bangladeshis living in America and comparisons with other populations. *Ergonomics*. 2009, 52, 8, s. 987–998. doi: 10.1080/00140130902792478.
- [26] JOHNSON, B. – OTIENO, W. – CAMPBELL-KYUREGHYAN, N. Influence of Jackhammer Weight on Grip Pressure, Muscle Activity, and Hand–Arm Vibration of the Operator. *IISE Transactions on Occupational Ergonomics and Human Factors*. 01 2017, 5, 1, s. 12–22. doi: 10.1080/24725838.2016.1269852.
- [27] KANCHAN, T. – KRISHAN, K. Anthropometry of hand in sex determination of dismembered remains-A review of literature. *Journal of Forensic and Legal Medicine*. 2011, 18, 1, s. 14–17. doi: 10.1016/j.jflm.2010.11.013.
- [28] KIPP, A. *Shinyapps.io - Getting started* [online]. may 2017. Dostupné z: <https://shiny.rstudio.com/articles/shinyapps.html>.
- [29] MANDAHAWI, N. et al. Hand anthropometry survey for the Jordanian population. *International Journal of Industrial Ergonomics*. 11 2008, 38, 11-12, s. 966–976. doi: 10.1016/j.ergon.2008.01.010.
- [30] MARR, B. How much data do we create every day? The mind-blowing stats everyone should read. *Forbes May*. 2018, 21, s. 2018.
- [31] MIRMOHAMMADI, S. J. et al. Anthropometric hand dimensions in a population of Iranian male workers in 2012. *International Journal of Occupational Safety and Ergonomics*. 12 2016, 22, 1, s. 125–130. doi: 10.1080/10803548.2015.1112108.
- [32] NAG, A. – NAG, P. – DESAI, H. Hand anthropometry of Indian women. *Indian Journal of Medical Research*. 2003, 117, s. 260–269.
- [33] NG, P. K. – SAPTARI, A. – MOHD FAUZI, A. *Hand Anthropometry: A Descriptive Analysis on Elderly Malaysians*, s. 187–201. 01 2013. doi: 10.13140/2.1.2302.3368. ISBN 978-967-12137-2-8.
- [34] OKUNRIBIDO, O. O. A survey of hand anthropometry of female rural farm workers in Ibadan, Western Nigeria. *Ergonomics*. 03 2000, 43, 2, s. 282–292. doi: 10.1080/001401300184611.
- [35] OVIEDO-TRESPALACIOS, O. et al. Hand anthropometric study in northern Colombia. *International Journal of Occupational Safety and Ergonomics*. 2017, 23, 4, s. 472–480. doi: 10.1080/10803548.2016.1217653.
- [36] PIERCE, C. A. – BLOCK, R. A. – AGUINIS, H. Cautionary note on reporting eta-squared values from multifactor ANOVA designs. *Educational and*

- psychological measurement*. 2004, 64, 6, s. 916–924. doi: 10.1177/0013164404264848.
- [37] REIF, J. *Metody matematické statistiky*. Západočeská univerzita, 2004. ISBN 80-7043-302-7.
- [38] ROUSSEEUW, P. J. – RUTS, I. – TUKEY, J. W. The bagplot: a bivariate boxplot. *The American Statistician*. 11 1999, 53, 4, s. 382–387. doi: 10.1080/00031305.1999.10474494.
- [39] RYCHTECKÁ, M. Smíme vědět, co nosíme? Zpráva z průzkumu mapující dostupnost informací o původu oděvů. *Společnost pro Fair Trade v rámci kampaně UŠILI TO NA NÁS! Kampaň za lepší pracovní podmínky při výrobě oděvů*. 2010.
- [40] TEAM. *shinyapps.io user guide* [online]. 04 2019. [cit. 2019/05/01]. Dostupné z: <https://docs.rstudio.com/shinyapps.io/>.
- [41] TRYBUS, M. et al. Hand anthropometry in patients with carpal tunnel syndrome: a case-control study with a matched control group of healthy volunteers. *Folia morphologica*. 05 2019, 78, 1, s. 182–190. doi: 10.5603/FM.a2018.0049.
- [42] TSAO, L. – MA, L. Using subject-specific three-dimensional (3D) anthropometry data in digital human modelling: case study in hand motion simulation. *Ergonomics*. 02 2016, 59, 11, s. 1526–1539. doi: 10.1080/00140139.2016.1151554.
- [43] MENSVOORT, M. *Elemental hand shape assessment 2.0: 'finger length vs. palm size proportions'!* [online]. 04 2017. [cit. 2019/04/28]. Dostupné z: <http://www.handresearch.com/news/elemental-hand-shapes-2-0-finger-length-vs-palm-size-proportions.htm>.
- [44] WANG, B. – SHI, W. – MIAO, Z. Confidence analysis of standard deviational ellipse and its extension into higher dimensional Euclidean space. *PloS one*. 03 2015, 10, 3, s. e0118537. doi: 10.1371/journal.pone.0118537.
- [45] WANG, C.-y. – CAI, D.-c. Hand tool handle design based on hand measurements. *MATEC Web of Conferences*. 01 2017, 119, s. 01044. doi: 10.1051/mateconf/201711901044.

Seznam obrázků

3.1	Popis bagplotu	9
3.2	Pravidlo 3-sigma, zdroj [44]	10
4.1	Náhled webové aplikace	22
4.2	Schéma webové aplikace	23
4.3	Metoda hlavních komponent - dataset iris setosa	29
5.1	Měřené rozměry ruky	32
5.2	Boxploty délky pravé a levé ruky v závislosti na pohlaví . . .	37
5.3	Funkce hustoty pro šířku pravé dlaně a pro obvod levé dlaně	38
5.4	Korelační matice a konfidenční interval pro pravý palec - ženy	40
5.5	Graf délky pravé a levé ruky s ohledem na pohlaví	40
5.6	Graf šířky a obvodu levé dlaně s ohledem na pohlaví	41
5.7	Graf závislosti BMI na obvodu pravé dlaně s ohledem na pohlaví	42
5.8	Boxploty šířky pravé i levé dlaně v závislosti na typu práce .	43
5.9	Boxploty obvodu pravé i levé dlaně v závislosti na dominanci ruky	44
5.10	Boxploty délky pravé ruky v závislosti na věkové skupině . .	44
5.11	Boxploty délky levé ruky v závislosti na BMI kategorii . . .	45
5.12	Paretův graf pro rozměry pravé ruky - typ povolání	46
5.13	Závislost první a druhé hlavní komponenty v závislosti na typu povolání	46
5.14	Závislost dat první a druhé hlavní komponenty se shluky v zá- vislosti na typu povolání	47
5.15	Závislost první a druhé hlavní komponenty se shluky v závis- losti na BMI kategoriích - levá ruka	47
5.16	Porovnání s populacemi - délka ruky	49
5.17	Porovnání s populacemi - délka dlaně	49
5.18	Porovnání s populacemi - šířka dlaně	50
5.19	Porovnání s populacemi - délka prostředníku	50
B.1	Uspaná aplikace na platformě shinyapps.io	64
B.2	Popis boxplotu	68
B.3	Histogram s funkcí hustoty, Q-Q graf pro normální rozdělení	68
B.4	Ukázka vhodně vytvořeného souboru pro porovnání dvou po- pulací	70

B.5 ANOVA - výstup v aplikaci	71
B.6 Kruskal-Wallisův test - výstup v aplikaci	71

Seznam tabulek

3.1	Schéma tabulky ANOVA	18
5.1	Přehled osob rozdělených do populačních skupin	35
5.2	Počty leváků a praváků	35
5.3	Počty osob rozdělených podle kategorie BMI	36
5.4	Počty osob rozdělených podle kategorie BMI	36
5.5	Korelační matice - pravá ruka	46
5.6	Rozptyl hlavních komponent	48
B.1	Přehled základních charakteristik antropometrických dat . .	77
B.2	Přehled základních charakteristik antropometrických dat 2 .	78
B.3	Přehled základních charakteristik antropometrických dat 3 .	79
B.4	Testy normality - délka ruky	80
B.5	Testy normality - délka dlaně	81
B.6	Testy normality - šířka dlaně	82
B.7	Testy normality - obvod dlaně	83
B.8	Testy normality - délka ukazováku	84
B.9	Testy normality - délka palce	85
B.10	Testy normality - délka pro stisk spouště	86
B.11	Testy normality - délka prostředníku	87
B.12	Korelační koeficienty - rozměry s výškou	88
B.13	Korelační koeficienty - rozměry mezi sebou	88
B.14	Parciální korelační koeficienty po vyřazení výšky - rozměry mezi sebou	89
B.15	Korelační koeficienty - porovnání s věkem	89
B.16	Parciální korelační koeficienty - porovnání s věkem	89
B.17	Korelační koeficienty - porovnání se šířkou	90
B.18	Parciální korelační koeficienty - porovnání se šířkou	90
B.19	Korelační koeficienty - porovnání s BMI	90
B.20	Parciální korelační koeficienty - porovnání s BMI	90
B.21	Korelační koeficienty - porovnání se šířkou (typ práce)	91
B.22	Korelační koeficienty - porovnání se šířkou (typ práce)	91
B.23	Korelační koeficienty - porovnání s BMI (typ práce)	91
B.24	Parciální korelační koeficienty - porovnání s BMI (typ práce)	92
B.25	ANOVA - pohlaví	92
B.26	ANOVA - dominantní ruka	92

B.27 ANOVA - typ práce	93
B.28 ANOVA - věková skupina	93
B.29 ANOVA - BMI kategorie	93
B.30 Přehled vybraných rozměrů ruky pro různé populace - ženy .	94
B.31 Přehled vybraných rozměrů ruky pro různé populace - muži	95

A Seznam použitých zkratek

Zkratka	Popis
F/M	pohlaví - žena/muž
L/R	dominance ruky - levák/pravák
K/M	typ práce - kancelářská/manuální
BMI	index tělesné hmotnosti
RSI	poškození z opakovaného namáhání, angl. <i>Repetitive Strain Injury</i>
CTD	poškození z opakovaného namáhání, angl. <i>Cumulative Trauma Disorder</i>
MAD	mediánová absolutní odchylka
PCA	metoda hlavních komponent
x	pozorování, proměnná
n	počet pozorování
k	počet kategorií
μ	střední hodnota
\bar{x}	aritmetický průměr
\tilde{x}, med	medián
\hat{x}	modus
R	variační rozpětí
x_{max}	maximální hodnota
x_{min}	minimální hodnota
σ^2	rozptyl
$s^2(x)$	výběrový rozptyl
$s(x)$	výběrová směrodatná odchylka
$\hat{\sigma}$	neparametrický odhad směrodatné odchylky
$s_{\bar{x}}(x)$	směrodatná chyba průměru
$s_{s(x)}(x)$	směrodatná chyba směrodatné odchylky
$s_p(x)$	směrodatná chyba 5. a 95. percentilů
VK	variační koeficient
x_q	kvantil s hladinou q
S	koeficient šikmosti
K	koeficient špičatosti
l	škálovací konstanta
$erfc$	komplexní chybová funkce
$F(x)$	distribuční funkce
$F_0(x)$	teoretická distribuční funkce

A Seznam použitých zkratek

r_{xy}	výběrový korelační koeficient
s_{xy}	výběrová kovariance
α	hladina významnosti testu
H_0	nulová hypotéza
H_1	alternativní hypotéza
SS_T	celkový součet čtverců variability
SS_A	součet čtverců variability mezi skupinami
SS_E	součet čtverců variability uvnitř skupin
MS_A	odhad rozptylu, průměrné čtverce variability mezi skupinami
MS_E	odhad rozptylu, průměrné čtverce variability uvnitř skupin
η^2	procentuální podíl rozptylu pro použití v ANOVA
F	F statistika
KW	Kruskall-Wallisova statistika

B Přílohy

B.1 Instalace R a RStudio

Tato část je věnována instalaci programovacího jazyka R a vývojového prostředí RStudio. Následující kroky ilustrují konkrétní postup pro operační systém Windows:

- Na adrese <https://cran.r-project.org/bin/windows/base/> je třeba stáhnout běhové prostředí jazyka R a to stiskem na odkaz *Download R x.x.x for Windows*, kde x.x.x je číslo poslední vydané verze, po stažení provést instalaci.
- Na adrese <https://www.rstudio.com/products/rstudio/download/> je třeba stáhnout vývojové prostředí RStudio a to stiskem na příslušný odkaz *RStudio x.x.x Windows 7+* v dolní části stránky, kde x.x.x je opět číslo poslední vydané verze, po stažení je nutné RStudio nainstalovat.

V případě instalace na jiný operační systém (OS) jsou kroky obdobné, jen je třeba ve výše uvedených bodech stáhnout variantu pro daný OS.

B.2 Spuštění aplikace na vlastním počítači

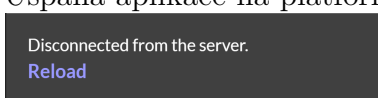
Pro případnou úpravu aplikace či její otestování na vlastním počítači je nezbytné nainstalovat knihovny, které aplikace používá, otevřít soubor `main.R` a aplikaci spustit. Následující kroky ilustrují konkrétní postup:

- Spustí se R Studio.
- Poté následuje otevření instalace `Balicku.R`.
- Stiskem kláves `CTRL+A` se provede výběr kódu souboru `instalaceBalicku.R`.
- Vybraný kód obsahující příkazy `install.packages()` se spustí stiskem tlačítka *Run* v pravém horním rohu okna.
- Je třeba vyčkat na instalaci balíčků (může trvat až několik minut v závislosti na rychlosti připojení a použitého hardware).
- Nyní lze soubor `instalaceBalicku.R` zavřít.
- Dále je nutné otevřít soubor `main.R`.
- Spuštění aplikace se provede stiskem tlačítka *Run App* v pravém horním rohu okna.

B.3 Uživatelská příručka

Webová aplikace je dostupná na stránce <https://datatesting.shinyapps.io/final/>. Pokud je uspaná (sleeping), stačí kliknout na ikonu *Reload* viz obr. B.1 Skládá se z hlavní nabídky, panelu možností na levé straně a části s hlavním obsahem, jež je rozdělen do několika sekcí pomocí záložek. Hlavní nabídka se skládá z krátkého popisu aplikace (Introduction), načtení dat (Data Loading), základních statistik (Basic Statistics), pokročilých statistik (Advanced Statistics) a porovnání s ostatními populacemi (Comparison). Veškeré popisky jsou v aplikaci napsané v anglickém jazyce.

Obrázek B.1: Uspaná aplikace na platformě shinyapps.io



Aplikace byla navržena tak, aby při nahrání libovolného souboru formátu .xlsx či .xls, který obsahuje kategorické i numerické proměnné, se provedly statistické testy. Výsledky ve formátu tabulek, grafů a výpisů se objeví ve výstupním panelu. Grafy lze uložit přes stisk pravého tlačítka myši a výběrem možnosti *Save Image*.

B.3.1 Příprava datového souboru

Nejprve je třeba zkontrolovat názvy sloupců a řádků. Obecně je vhodné nastavit první řádek jako hlavičku tabulky s pojmenováním. První sloupec potom obsahuje jednotlivá pozorování, které je pojmenováno unikátním výrazem (obvykle číselnou hodnotou od 1 do počtu pozorování). Pokud se některé názvy shodují, název je automaticky pozměněn s číslem na konci (př. "minimum, minimum"-> "minimum...1, minimum...2"). [4]

Jazyk R je citlivý na velikost písmen (case sensitive). Pojmenování by mělo být konzistentní. Speciální symboly jsou povoleny (např. "?,\$,*,+,-,/,({,|,>,"=). Název může začínat i číslicí.

Prázdné údaje by měly být smazány, rovněž chybějící hodnoty nahrazeny textem "NA" nebo hodnotou 9999. Údaje obsahující datum by měly být ve formátu 4-místného roku: "yyyy-mm-dd".

Takto upravený soubor je možné nahrát do aplikace, ovšem není to podmínkou. Je důležité poznamenat, že je vhodné veškeré pojmenování provádět v anglickém jazyce, neboť ne všechna interpunkční znaménka se podaří uchovat. Také je vhodné, aby soubor obsahoval kategorické proměnné, neboť určité tabulky a grafy jsou vykresleny v závislosti na vybrané kategorii.

B.3.2 Načtení dat (sekce Data Loading)

Do aplikace je možné nahrát pouze soubory aplikace Excel zakončené příponou *.xlsx* či *.xls*. V panelu možností si uživatel pomocí tlačítka *Browse* vybere soubor, vyplní jméno listu načítaného souboru (výchozí hodnota je *List1*) a stiskne tlačítko *Load data*. Dle velikosti souboru potrvá i načtení souboru (do velikosti 1000 řádků řádově vteřiny). Pokud je soubor načten správně, objeví se tabulka s nahranými daty z vybraného listu.

Stane-li se, že uživatel zadá název listu souboru, který neexistuje, aplikace ukončí svůj běh. V tomto případě stačí stránku s aplikací znovu načíst. Ovšem pokud nahraje soubor správně a napodruhé zadá špatný název listu, uživatel je upozorněn chybovou hláškou "Error: Cannot find sheet named "ListX".

Během načítání souboru se provádí kontrola, která maže řádky, jejichž hodnoty neleží ve vzdálenosti tří směrodatných odchylek od mediánu (směrodatná odchylka je odhadnuta metodou MAD). Proto, ne všechna data jsou nahrána. Přehled řádků (id probandů), které byly smazány, jsou viditelné v sekci *Data Loading*. Na další sekci se lze dostat přes stisknutí tlačítka *Move to Basic Statistics* nebo překliknutím na další panel.

B.3.3 Základní statistiky (sekce Basic Statistics)

V panelu možností lze vybrat proměnné z načtených dat. Pokud data nejsou nahrána, objeví se popisek *No data loaded*. První rozbalovací seznam (*Choose a variable*) obsahuje všechny proměnné (numerické i kategorické). Vybere-li uživatel numerickou proměnnou, po stisku tlačítka *Load Basic Characteristics* se ukáže tabulka četností, kvantilů a popisných charakteristik v záložce *Table* a koláčový graf, histogram a odhad funkce hustoty v záložce *Plot*. V případě kategorické proměnné se objeví tabulka četností, koláčový a sloupcový graf.

Dalšími dvěma vstupy jsou výběr numerických proměnných X a Y (*Choose a X variable for X-Y analysis (numerical)*, *Choose a Y variable for X-Y analysis (numerical)*) a kategorie (*Choose a category for X-Y analysis*). Po stisku tlačítka *Load basic characteristics* se v záložce *X-Y analysis* objeví popisné charakteristiky pro proměnnou X a Y zvlášť a vykreslí se grafy boxplotů a bagplotů podle dané kategorie. Na záložce *Normality of X-Y variables* jsou k dispozici výsledky testů normality (Jarque-Bera test, Lilieforsův test, Sharipův test, Kolmogorovův-Smirnovův test), koeficienty šikmosti a špičatosti, pro grafické znázornění histogramy, grafy funkcí hustoty a Q-Q grafy pro proměnnou X i Y zvlášť v závislosti na vybrané kategorii.

Na další záložce *Correlation Analysis* jsou k dispozici korelační koeficienty, parciální korelační koeficienty a příslušné p-hodnoty testů významnosti těchto koeficientů. Vzhledem k tomu, že parciální korelační koeficienty zkoumají vztahy 3 proměnných, v rozbalovacím seznamu je k dispozici další výběr proměnné označen *Choose a Z variable for partial correlation coefficient (numerical)*. Dále na této záložce lze vidět bodový graf s přímkou sestrojenou dle metody nejmenších čtverců a s 95% konfidenčním intervalem spolehlivosti a to na základě vybrané kategorie. Posledními grafy na této záložce jsou grafy korelačního Pearsonova koeficientu a korelační matice, jež mají tmavou barvu v případě pozitivní korelace a červenou barvu v případě negativní korelace.

Poslední záložkou této sekce je přehled (*Summary*, který ukáže základní deskriptivní charakteristiky a percentily pro danou kategorii. Nicméně toto členění se ještě dělí podle další skupiny, kterou je třeba vybrat ve výběrovém panelu. Jedná se o možnost *Choose a second category for summary*. Lze tak zobrazit například kolik mužů a žen pracuje manuálně a kolik v kanceláři.

Doplňující informace

Stručný popis popisných charakteristik [23], které nebyly uvedeny v kapitole 3.2.1, jsou uvedeny níže. Rovněž tak popis boxplotu.

Aritmetický průměr, značený \bar{x} , je definován jako součet všech pozorování vydělený jejich počtem. Je tedy jakousi střední hodnotou daného souboru, a proto se při odhadu střední hodnoty náhodného rozdělení pravděpodobnosti používá právě aritmetický průměr.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}, \quad (\text{B.1})$$

n ... počet pozorování

x ... pozorování, $x \in R$.

Ovšem ne vždy aritmetický průměr ukazuje to, co je požadováno. V případě součtu záporných a kladných hodnot se určitá informace ztratí. Proto je někdy vhodné se zaměřit na hodnoty modusu a mediánu. **Medián**, značeno \tilde{x} , je hodnota dělící seřazenou posloupnost hodnot na dvě stejně dlouhé poloviny. Používá se, pokud je chtěno znát střed rozdělení dat a pomohou nám při identifikaci odlehlých pozorování. Při výpočtu se rozlišuje, zda počet pozorování n je sudé či liché. Pro n sudé

$$\tilde{x} = \frac{x_{n/2} + x_{n/2+1}}{2}, \quad (\text{B.2})$$

pro n liché

$$\tilde{x} = \frac{x_{(n+1)/2}}{2}. \quad (\text{B.3})$$

Modus, značený \hat{x} , označuje hodnotu s největší četností, tudíž hodnotu, která se v dané proměnné nejčastěji opakuje.

Variační rozpětí, značeno R , se počítá jako rozdíl minimální od maximální hodnoty. Říká, jak velká je rozptýlenost dat. V případě odlehlých hodnot je tato hodnota charakteristiky velká, tudíž může zkreslit výsledek. Je tedy citlivá na odlehlá pozorování.

$$R = x_{max} - x_{min} \quad (\text{B.4})$$

Jak moc jsou data rozptýlena kolem aritmetického průměru udávají charakteristiky rozptylu a směrodatné odchylky. **Výběrový rozptyl**, značený $s^2(x)$, je definován jako součet kvadrátů odchylek od aritmetického průměru.

$$s^2(x) = \frac{\sum (x_i - \bar{x})^2}{n - 1} \quad (\text{B.5})$$

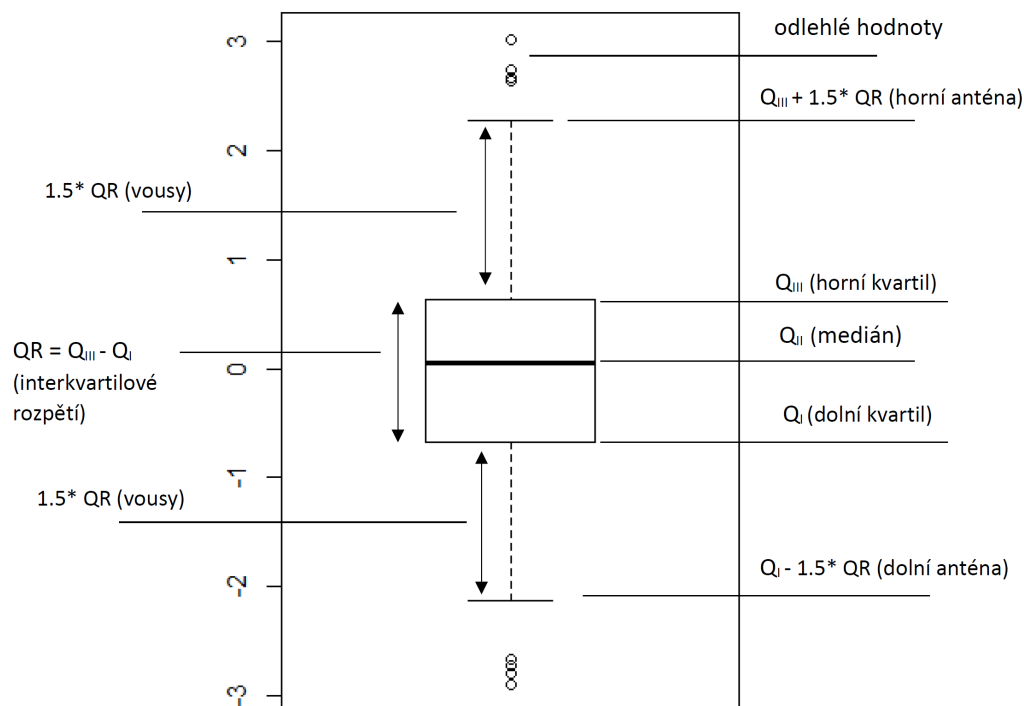
Výběrová směrodatná odchylka, značena $s(x)$, je pak odmocnina z rozptylu. Používá se pro zjištění míry rozptýlenosti dat kolem průměru. Směrodatná odchylka je buďto nulová, v případě shodnosti všech dat, či je kladná. Ovšem i tato charakteristika je citlivá vůči odlehlým hodnotám.

$$s(x) = \sqrt{s^2(x)} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}} \quad (\text{B.6})$$

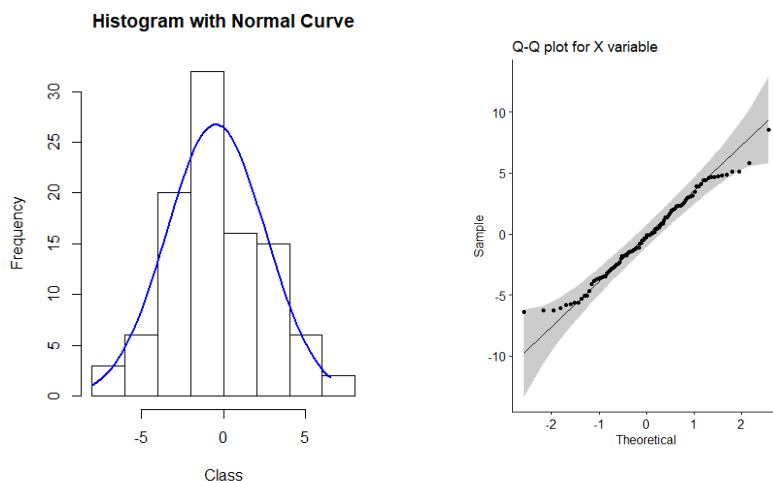
Boxplot Krabicový diagram neboli boxplot se skládá z krabičky, která zobrazuje medián, horní a dolní kvartil a tím tedy interkvartilové rozpětí (rozdíl dolního a horního kvartilu), jeho vousků vyjadřující variabilitu dat a hodnot nad a pod vousky, odlehlých pozorování. Jeho graf s popisky je znázorněn na obr. B.3.3.

Histogram je sloupcový graf, kde počet sloupců reprezentuje počet tříd a výška sloupců danou četnost. Počet tříd (k) je stanoven podle Sturgesova pravidla $k = (1 + 3, 32 \log(n))$. Funkce hustoty normálního rozdělení má tvar Gaussovy křivky, histogram s funkcí hustoty na vygenerovaných normálních datech s $n = 100$, $\mu = 0$ a $\sigma = 3$ lze vidět v grafu B.3. Dále lze vykreslit **Q-Q graf** (anglicky quantile-quantile plot). Tento graf řadíme mezi pravděpodobnostní graf, kde jsou porovnávány kvantily teoretického rozdělení s naměřenými kvantily. Vykreslenými body se proloží přímka pomocí metody nejmenších čtverců. Leží-li body okolo přímky v konfidenčním intervalu, jedná se o normální rozdělení, neboť teoretické a naměřené kvantily se od sebe příliš neliší.

Obrázek B.2: Popis boxplotu



Obrázek B.3: Histogram s funkcí hustoty, Q-Q graf pro normální rozdělení



B.3.4 Pokročilá statistika (sekce Advanced Statistics)

Další sekci se rozumí statistické metody, které nepatří mezi základní charakteristiky. V našem případě se jedná o analýzu rozptylu (ANOVA), metodu hlavních komponent (PCA) a shlukování. Nalevo v panelu možností je třeba zaškrtnout všechny proměnné, které chtějí být v této části zkoumány

(nejméně dvě). Postupně je lze měnit. Je třeba vybrat rovněž kategorie, podle kterých budou testovány hypotézy (v případě analýzy rozptylu, zda existuje statisticky významný rozdíl mezi vybranými kategoriemi). Pro výsledky je nutno stisknout tlačítko *Load Advanced Characteristics*.

První záložka *Summary of chosen data* dává krátký přehled vybraných statistik (průměr, směrodatné odchylnka, medián, 5. a 95. percentil) vybraných proměnných.

V záložce *ANOVA* je k dispozici kompletní přehled testů týkající se analýzy rozptylu. Jako první se zobrazí vykreslené boxploty, jež umožňují vidět přehled mezi danými skupinami pro každou proměnnou zvlášť. Dále jsou kontrolovány předpoklady (normalita, shodnost rozptylů dat). K tomu je využito Jarque-Berova testu normality, F-testu (v případě, že kategorie není větší než 2) či Levenova testu. Je rozlišováno, kolik kategorií má vybraná skupina. Pokud má skupina maximálně 2 kategorie, provádí se Welchův t-test a Wilcoxonův test. Pokud je skupina složena z více kategorií, využívá se jednofaktorové analýzy rozptylu (je vypisována i informace o procentuálním vyjádření rozptylu jež vysvětluje nezávislá proměnná), a dále Kruskal-Wallisova testu. Párové porovnání mezi skupinami se provádí pomocí Tukeyho testu (v případě jednofaktorové analýzy rozptylu) a párového Wilcoxonova testu (v případě neparametrické obdoby jednofaktorové analýzy rozptylu). Výsledky parametrických i neparametrických testů se vypisují v aplikaci jako výstup z konzole (viz obr. B.6, B.5).

Poslední záložka *PCA and Clustering* obsahuje kontrolu předpokladů použití metody hlavních komponent, korelační matici, Paretův graf, procentní podíl vysvětlených rozptylů pro hlavních komponent a vykreslení hlavních komponent s informací o shlucích pro dvě hlavní komponenty. Konfidenční množiny jsou rovněž vykresleny (t-rozdělení plně a normálního rozdělení čárkovaně).

B.3.5 Porovnání s ostatními populacemi (sekce Comparison)

V této sekci se provádí porovnání s ostatními populacemi. Proto je nejprve třeba připravit datový soubor, který v tomto případě musí mít danou strukturu. Prvním sloupcem je název státu/národnosti, druhým pohlaví, třetím počet pozorování, čtvrtý a další sloupec je vyhrazen pro rozměry dané populace. Uvedou se vždy informace o průměru a směrodatné odchylce v tomto pořadí (jako např. v obr. B.4).

Obrázek B.4: Ukázka vhodně vytvořeného souboru pro porovnání dvou populací

Data Loading		Comparison with Other Nationalities						
	States	Sex	Number of Observations	Hand breadth with metacarpals (šířka.dlaně)	X_1	Hand length (délka ruky dlaňová)	X_2	
1	Americans (USA)	F	211.00					
2	Bangladeshi (USA)	F	51.00	74.12	4.05	167.64	11.49	
3	Bangladeshi (USA)	M	50.00	79.39	7.16	173.15	8.38	
4	Brazilian	F	206.00	78.17	3.83	177.30	8.91	
5	Brazilian	M	2123.00	87.70	4.27	190.40	10.35	
6	British	F	92.00					
7	Colombian	F	126.00	78.80	3.40	167.30	8.00	
8	Colombian (northern)	M	129.00	82.65	9.42	190.34	13.60	
9	Czechs	F	433.00	78.80	4.04	176.02	7.91	
10	Czechs	M	554.00	88.39	5.37	191.99	9.43	
11	Filipino	F	962.00	92.30	6.97	179.50	3.44	

Uživatel načte soubor pomocí tlačítka *Browse*, vybere list (*Choose sheet in Excel*) a stiskem tlačítka *Load data* načte soubor. Uživatel může editovat nahraný soubor a poté stiskem tlačítka *Save* uloží soubor v aplikaci (ne do počítače). Z rozbalovacího seznamu (*Select a Variable*) vybere rozměr a zaškrtně kategorie, jež chce porovnávat (muže M, ženy F, oba dohromady). Opět pokud uživatel nezaškrtně, je upozorněn chybovou hláškou. Vzhledem k tomu, že kdokoliv může nahrát soubor a porovnávat svou národnost, v panelu výběrů je ještě jeden výběr *Select the Main Nationality*. Uživatel zaškrtně název národnosti, se kterou chce provést porovnání. Poté je stiskem tlačítka *Compare* provedeno porovnání dat. V závislosti na kategorii dat a výběru národnosti pro porovnání se objeví tabulky rozdílů (od národnosti je odečítána vybraná národnost; je-li tedy rozdíl kladný, znamená to, že vybraná národnost k porovnání má větší hodnoty než porovnávaná), dále jsou graficky znázorněny 95% intervaly spolehlivosti pro každou národnost v závislosti na vybrané kategorii. Jsou uvedeny i t-statistiky Welchova testu pro nestejně rozptýly, dolní a horní hranice intervalu a p-hodnoty.

Obrázek B.5: ANOVA - výstup v aplikaci

One-way ANOVA

```
[1] "Right thumb length (délka pravého palce-6)"
[1] "ANOVA 1factorial"
[1] "Assumptions are fulfilled (normal distribution, homoscedasticity)."
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	5	903	180.55	8.422	8.22e-08 ***
Residuals	968	20751	21.44		

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
[1] "ANOVA1: There are significant differences between groups, the p_value is 8.22216971694956e-08 ."
[1] "Tukey Honestly Significant Difference between means of groups:"
  Tukey multiple comparisons of means
  95% family-wise confidence level

Fit: aov(formula = serazenaData[, j] ~ group, data = serazenaData)

$group
      diff      lwr      upr    p adj
30-39-20-29  2.0334736  0.7577997  3.3091474  0.0000876
40-49-20-29  1.8758140  0.6791196  3.0725085  0.0001239
50-59-20-29  2.2240560  0.9731859  3.4749261  0.0000068
60-69-20-29  0.6758461 -1.0851563  2.4368485  0.8831033
70-79-20-29  1.6642713 -0.3505371  3.6790798  0.1720823
40-49-30-39 -0.1576595 -1.5980423  1.2827232  0.9996030
50-59-30-39  0.1905824 -1.2951164  1.6762812  0.9991398
60-69-30-39 -1.3576274 -3.2925058  0.5772509  0.3409424
70-79-30-39 -0.3692022 -2.5376298  1.7992253  0.9966504
50-59-40-49  0.3482420 -1.0702201  1.7667041  0.9818260
60-69-40-49 -1.1999679 -3.0837111  0.6837752  0.4539191
70-79-40-49 -0.2115427 -2.3344680  1.9113826  0.9997494
60-69-50-59 -1.5482099 -3.4668257  0.3704059  0.1932559
70-79-50-59 -0.5597847 -2.7137137  1.5941444  0.9765987
70-79-60-69  0.9884252 -1.4967033  3.4735537  0.8664414

[1] "4.2 percent of variance is explained by the independent variable."
```

Obrázek B.6: Kruskal-Wallisův test - výstup v aplikaci

Non-parametric alternative of one-way ANOVA: Kruskal-Wallis test

```
[1] "Right hand length (délka pravé ruky dlaňová-1)"
[1] "Kruskal-Wallis test (non-parametric alternative to one-way ANOVA):"
```

Kruskal-Wallis rank sum test

```
data: serazenaData[, j] by group
Kruskal-Wallis chi-squared = 22.26, df = 4, p-value = 0.0001779
```

```
[1] "Kruskal-Wallis: There are significant differences between groups, the p_value is 0.000177862789312376 ."
[1] "Pairwise comparisons: Differences between pairs of groups:"
```

Pairwise comparisons using Wilcoxon rank sum test

```
data: serazenaData[, j] and group
```

	ideální váha	mírná obezita	nadváha	podváha
mírná obezita	1.0000	-	-	-
nadváha	0.0547	0.1639	-	-
podváha	0.0013	0.0012	0.0155	-
střední obezita	1.0000	1.0000	1.0000	0.2255

```
P value adjustment method: holm
```

B.4 Přílohy na CD

Na přiloženém CD jsou obsaženy následující vytvořené soubory. CD obsahuje složky Aplikace, Data, Tex a Poster. Složka Aplikace obsahuje všechny zdrojové soubory programového řešení (7 souborů). Složka data, obsahuje původní data, upravená data a data pro testování (6 souborů). Složka Tex obsahuje zdrojové text práce s obrázky. Ostatní soubory jsou v kořenovém adresáři CD. Jedná se o tisknutelnou podobu práce ve formátu .pdf (DP_Koukalova.pdf) a soubor aplikace Excel se shrnutými výsledky z aplikace (Tabulky_Porovnání_Rozměry.xlsx). Ve složce Poster jsou k dispozici soubory Koukalova_Jana_2019.pub i Koukalova_Jana_2019.pdf.

Složka Aplikace obsahuje:

- main.R - hlavní kód webové aplikace, obsahuje UI i serverovou část, pro spuštění aplikace
- predzpracovani_sigma.R - funkce pro detekci odlehlých pozorování pomocí metody 3-sigma, odstranění neplatných hodnot, úprava kategoričkových proměnných
- predzpracovani_MAD.R - funkce pro detekci odlehlých pozorování pomocí metody MAD odstranění neplatných hodnot, úprava kategoričkových proměnných
- zakladniCharakteristiky.R - zpracování deskriptivních charakteristik, testy normality, korelační analýza, přehledy
- pokrocilaAnalyza.R - ANOVA, PCA, shlukování
- porovnaní.R - zpracování výsledků pro porovnání
- instalaceKnihoven.R - skript pro instalaci potřebných knihoven

Složka Data obsahuje:

- Data_Původní.xlsx - původní data o antropometrii rukou získaná od RTI ZČU
- Data_Upravená.xlsx - upravená data pro analýzu
- Data_Porovnání.xlsx - vytvořená databáze antropometrie rukou pro různé populace
- data_outliers.xlsx - vytvořený dataset pro identifikaci odlehlých pozorování k testování
- data_na.xls - vytvořený dataset pro identifikaci neplatných hodnot k testování
- iris_test.xlsx - volně dostupný dataset iris

B.5 Tištěné přílohy

Listing B.1: Předzpracování

```

predzpracovani_2 <- function(puvData){
# detekce odlehlych pozorovani, pravidlo MAD
# kontrola na neplatne hodnoty – prazdne hodnoty, 9999

odstranenaIdNeplatneHodnoty<-c()
odstranenaIdOdhlehlaPozorovani<-c()

#detekce a odstraneni radku s~neplatnymi hodnotami
for (sloupec in 2:ncol(puvData)){
  if (is.numeric(puvData[,sloupec])) {
    #kontrola na numericke promenne
    indexyRadku<-c()
    for (radek in 1:nrow(puvData)) {
      if (is.na(puvData[radek, sloupec]) |
        puvData[radek, sloupec] == "" |
        puvData[radek, sloupec] == 9999) {
        indexyRadku<-c(indexyRadku, radek)
        odstranenaIdNeplatneHodnoty<-
c(odstranenaIdNeplatneHodnoty, puvData[radek, 1])
      }
    }

    if (is.null(indexyRadku)==FALSE) { #je prazdne pole
      for (i in 1:length(indexyRadku)){
        puvData<-puvData[-(indexyRadku[i]-(i-1)),]
        #smazani radku
      }
    }
  }

#detekce a odstraneni radku
#kde numericcka promenna obsahuje odlehle pozorovani
pocetMAD <- 3 # nasobek MAD pro detekci odlehlych pozorovani
k<- -1/(sqrt(2)*(qnorm((3/2)/2, lower = FALSE)/sqrt(2)))
#konstanta pro metodu MAD

for (sloupec in 2:ncol(puvData)){
  if (is.numeric(puvData[,sloupec])) { #numericcka promenna
    numerical<-puvData[,sloupec]
  }
}

```

```

indexyRadku<-c()

#detekce odlehlych pozorovani pomoci metody MAD
MAD <- k~* median(abs(numerical-median(numerical)))

medianDatSloupce <- median(numerical)

dolni <- medianDatSloupce - (pocetMAD * MAD)
horni <- medianDatSloupce + (pocetMAD * MAD)

for (radek2 in 1:nrow(puvData)){
  if (numerical[radek2] >=dolni &
      numerical[radek2]<=horni) {
    #lezi v intervalu +- 3 sigma
  } else {
    indexyRadku<-c(indexyRadku , radek2)
    odstranenaIdOdlehlaPozorovani<-
    c(odstranenaIdOdlehlaPozorovani ,
      puvData[radek2 ,1])
  }
}

if (is.null(indexyRadku)==FALSE) { #je prazdne pole
  for (i in 1:length(indexyRadku)){
    puvData<-puvData[-(indexyRadku[i]-(i-1)),]
    #smazani radku
  }
}

} else { #kategoricka promenna
  #odstraneni mezer pred a za textem
  puvData[, sloupec]<-gsub("^\\s+|\\s+$",
    "",
    puvData[, sloupec])
}
}
return(list(puvData ,
  sort(odstranenaIdOdlehlaPozorovani) ,
  sort(odstranenaIdNeplatneHodnoty)))
#return vraci vzdy jednu hodnotu, proto list
}

```

Listing B.2: Ukázka vykreslení bodového grafu

```

output$scatterplotXY <- renderPlot({
  cond=data[,input$cat1]

  tabulka<-data.frame(cond=cond,
    xvar=data[,input$x],
    yvar=data[,input$y])

  ggplot(data = tabulka,
    aes_string(x = tabulka$xvar,
               y = tabulka$yvar,
               color=tabulka$cond)) +
    geom_point()+
    geom_smooth(method=lm, se=TRUE)+
    #prolozeni primkou + confidence interval
    ggtitle("Scatterplot of X,Y...")+
    xlab(input$x)+
    ylab(input$y)+
    theme(plot.title=element_text(
      lineheight=.8, face="bold"))
})

```

Listing B.3: Ukázka vytvoření tabulky základních charakteristik

```

output$s3_content <- renderTable({
  #fce modus - je to cislo s~nejvetsi cetnosti
  modus<-which.max(tabulate(match(numerical,unique(numerical))))
  #variacni koeficient
  varKoeff<-(sd(numerical)/mean(numerical))*100
  #smerodatna chyba prumeru
  sdErrorMean<-sqrt(sd(numerical)/length(numerical))
  #smerodatna chyba smerodatne odchylky
  sdErrorSd<-sqrt(sd(numerical)/(2*length(numerical)))
  #smerodatna chyba 5 a 95 percentilu
  sdErrorPerc<-sqrt(sdErrorMean^2+(1.645^2*sdErrorSd^2))
  #variacni rozpeti
  range<-max(numerical)-min(numerical)

```

```

tab<-matrix(c(length(numerical),
               round(min(numerical),3),
               round(max(numerical),3),
               round(range,3),
               round(mean(numerical),3),
               round(var(numerical),3),
               round(sd(numerical),3),
               round(sdErrorMean,3),
               round(sdErrorPerc,3),
               round(varKoeff,3),
               modus,
               median(numerical)),
            ncol=1)

rownames(tab) <-c("Count", "Minimum", "Maximum",
                 'Range', "Mean", "Variance", 'Standard Deviation',
                 'Standard Error of the Mean',
                 "Standard Error of 5th & 95th percentile",
                 'Variance Coefficient',
                 'Mode', 'Median')
colnames(tab)<-"Value"
tab<-data.frame(tab)
}, striped=TRUE, hover=TRUE, bordered=TRUE,
rownames = TRUE, digits=3, colnames = TRUE,
align='c', caption = paste("Basic statistics", jmenoSloupce),
caption.placement = getOption("xtable.caption.placement",
                              "top"))

```

Tabulka B.1: Přehled základních charakteristik antropometrických dat

Rozměr	Skupina	Kat.	Pravá ruka (náměry)			Levá ruka (náměry)		
			\bar{x}	\bar{x}	s(x)	\bar{x}	\bar{x}	s(x)
Délka ruky [mm] (1)	Pohlaví	F	176,250	176,0	7,918	175,877	175,0	7,800
		M	192,210	192,0	9,430	191,522	191,0	9,250
	Dominance	L	186,476	187,5	11,368	186,952	187,5	11,147
		R	185,004	185,0	11,880	184,360	184,0	11,643
	Typ práce	K	184,360	184,0	12,767	183,984	184,0	12,767
		M	185,781	186,0	10,967	185,087	185,0	10,543
Délka dlaně [mm] (2)	Pohlaví	F	100,558	100,0	4,767	100,287	100,0	4,724
		M	110,133	110,0	5,853	109,561	110,0	5,855
	Dominance	L	106,333	107,0	6,683	106,560	108,0	6,487
		R	105,844	106,0	7,244	105,343	105,0	7,133
	Typ práce	K	105,544	105,0	7,635	105,202	105,0	7,689
		M	106,174	106,0	6,797	105,654	106,0	6,533
Šířka dlaně [mm] (3)	Pohlaví	F	78,412	78,0	4,040	79,185	79,0	4,045
		M	88,203	88,0	5,192	88,662	89,0	5,413
	Dominance	L	85,583	85,5	6,937	84,583	84,5	6,754
		R	83,698	83,5	6,741	84,447	84,0	6,767
	Typ práce	K	82,757	82,0	6,935	83,553	83,0	6,972
		M	84,788	85,0	6,501	85,221	85,0	6,491
Obvod dlaně [mm] (4)	Pohlaví	F	190,725	190,5	9,264	194,213	194,0	9,331
		M	216,190	215,0	11,627	220,090	220,0	11,710
	Dominance	L	209,107	211,5	17,948	209,202	211,0	17,231
		R	204,498	205,0	16,350	208,557	208,0	16,703
	Typ práce	K	202,906	203,0	17,016	206,796	206,0	17,331
		M	206,569	208,0	15,944	210,142	211,0	16,085
Délka ukazováku [mm] (5)	Pohlaví	F	68,919	69,0	4,032	68,662	69,0	4,042
		M	73,902	74,0	4,808	73,260	73,0	4,478
	Dominance	L	72,179	72,5	5,260	71,357	72,0	4,859
		R	71,646	72,0	5,105	71,208	71,0	4,862
	Typ práce	K	71,351	71,0	5,389	70,937	71,0	5,174
		M	71,979	72,0	4,865	71,459	71,0	4,569
Délka palce [mm] (6)	Pohlaví	F	58,627	59,0	3,972	60,113	60,0	3,834
		M	62,528	63,0	4,460	63,958	64,0	4,664
	Dominance	L	61,262	61,0	5,366	62,560	62,0	5,168
		R	60,754	61,0	4,599	62,224	62,0	4,675
	Typ práce	K	60,369	60,0	4,659	61,465	61,0	4,915
		M	61,159	61,0	4,653	62,915	63,0	4,443
Délka pro stisk spouště [mm] (7)	Pohlaví	F	66,595	66,0	5,052	67,252	67,0	4,655
		M	73,854	74,0	6,124	73,284	73,0	6,033
	Dominance	L	70,512	70,5	6,751	72,274	72,0	5,975
		R	70,646	70,0	6,723	70,452	70,0	6,235
	Typ práce	K	70,355	70,0	6,804	70,324	70,0	6,368
		M	70,870	71,0	6,649	70,849	71,0	6,109
Délka prostředníku [mm] (8)	Pohlaví	F	75,692	76,0	4,801	75,590	76,0	4,643
		M	82,077	82,0	5,180	81,961	82,0	4,990
	Dominance	L	80,143	80,0	5,685	80,393	80,5	5,917
		R	79,161	79,0	5,952	79,017	79,0	5,757
	Typ práce	K	78,816	79,0	6,436	78,782	79,0	6,188
		M	79,607	80,0	5,455	79,433	79,0	5,403

Tabulka B.2: Přehled základních charakteristik antropometrických dat 2

Rozměr	Skupina	Kategorie	Levá ruka (náměry)			Pravá ruka (náměry)		
			Průměr	Medián	s(x)	Průměr	Medián	s(x)
Délka ruky [mm] (1)	Věk,skup,	20-29	186,749	188	12,258	186,2	186	11,829
		30-39	187,188	187	11,578	186,766	186	11,722
		40-49	185,425	185	11,179	184,984	184	10,912
		50-59	183,785	183	11,587	183,074	182	11,084
		60-69	179,642	178	10,98	179,358	178	11,162
		70-79	177,816	177	8,301	176,653	177	8,731
	BMI skup,	ideální	183,376	184	11,645	182,93	183	11,19
		mírná ob,	186,65	187	10,644	186,049	185	10,838
		nadváha	187,345	187	12,168	186,757	187	12,048
		podváha střed,ob,	177,353	178	10,782	176,529	178	11,587
			186,067	187	8,98	183,933	184	10,06
Délka dlaně [mm] (2)	Věk,skup,	20-29	106,448	106	7,409	106,158	106	7,226
		30-39	107,312	108	7,186	107,175	107	7,124
		40-49	105,952	106	6,842	105,597	105	6,801
		50-59	105,233	105	7,138	104,374	104	6,892
		60-69	103,463	102	6,774	102,552	102	6,484
		70-79	102,571	102	5,891	101,837	102	5,647
	BMI skup,	ideální	104,371	104	7,017	104,209	104	6,831
		mírná ob,	107,415	108	6,474	106,585	107	6,503
		nadváha	107,574	108	7,26	106,913	107	7,35
		podváha střed,ob,	101,118	99	5,146	101	101	5,534
			107,8	107	6,625	106,6	106	6,277
Šířka dlaně [mm] (3)	Věk,skup,	20-29	83,358	85	6,972	83,772	84	7,013
		30-39	85,221	86	6,277	86,045	87	6,287
		40-49	84,694	83	7,055	85,317	84	6,98
		50-59	83,988	83	6,698	84,294	84	6,75
		60-69	82,612	82	6,665	83,657	83	6,591
		70-79	81,347	81	4,794	82,837	82	4,346
	BMI skup,	ideální	81,77	81	6,096	82,281	82	6,105
		mírná ob,	87,805	88	6,578	88,073	88	6,42
		nadváha	85,522	86	6,521	86,322	87	6,487
		podváha střed,ob,	77,059	76	5,932	78,353	77	6,441
			87,067	87	7,255	87,733	87	7,787
Obvod dlaně [mm] (4)	Věk,skup,	20-29	204,437	208	16,704	208,056	211	16,638
		30-39	207,727	210	15,915	211,76	214	16,306
		40-49	206,398	205	17,619	210,156	207,5	18,057
		50-59	205,025	204	16,049	208,73	206	15,86
		60-69	201,224	198	16,455	204,567	200	17,786
		70-79	198,204	195	11,761	202,041	198	11,876
	BMI skup,	ideální	199,253	198,5	14,91	202,996	201	15,095
		mírná ob,	215,293	216	15,92	218,236	218	16,17
		nadváha	209,383	210	15,173	213,339	215	15,588
		podváha střed,ob,	187,235	184	15,409	190,824	188	14,492
			214,733	215	18,649	218,667	221	19,107

Tabulka B.3: Přehled základních charakteristik antropometrických dat 3

Rozměr	Skupina	Kategorie	Levá ruka (náměry)			Pravá ruka (náměry)		
			Průměr	Medián	s(x)	Průměr	Medián	s(x)
Délka ukazováku [mm] (5)	Věk,skup,	20-29	71,938	72	5,383	71,487	72	4,966
		30-39	72,662	73	5,177	72,058	72	4,855
		40-49	71,941	71,5	4,637	71,441	71	4,661
		50-59	71,27	71	5,153	70,675	71	4,944
		60-69	70,179	70	4,482	70,06	70	4,667
		70-79	69,388	69	4,334	69,224	69	3,938
	BMI skup,	ideální	71,289	71	4,997	70,895	71	4,705
		mírná ob,	72,098	72	4,969	71,715	72	4,522
		nadváha	72,258	72	5,305	71,629	71	5,106
		podváha střed,ob,	69,353	70	5,111	69,059	70	5,607
			70,733	70	4,217	70,533	70	4,719
Délka palce [mm] (6)	Věk,skup,	20-29	60,617	60	4,967	61,648	61	5,114
		30-39	61,247	61	4,797	63,221	63	4,529
		40-49	61,016	61	4,201	62,538	62	4,094
		50-59	60,479	61	4,61	62,393	62	4,61
		60-69	60,507	61	4,436	62,358	62	4,562
		70-79	61,327	62	4,254	61,898	62	4,66
	BMI skup,	ideální	60,325	60	4,794	61,692	61	4,77
		mírná ob,	61,325	61	4,542	63,008	63	4,5
		nadváha	61,351	62	4,53	62,864	63	4,641
		podváha střed,ob,	58,235	58	4,409	59,529	59	4,584
			61,6	62	2,558	62,8	63	3,726
Délka pro stisk spouště [mm] (7)	Věk,skup,	20-29	71,665	72	7,065	71,38	72	6,651
		30-39	71,429	71	6,434	71,916	72	5,872
		40-49	70,817	71	6,236	70,613	70	5,399
		50-59	69,27	69	6,694	69,509	69	6,464
		60-69	68,716	68	6,132	68,418	67	5,713
		70-79	67,143	66	5,397	67,551	66	4,726
	BMI skup,	ideální	69,686	69	6,519	69,857	69	6,161
		mírná ob,	71,789	72	6,656	71,545	71	5,83
		nadváha	71,617	72	6,802	71,397	71	6,214
		podváha střed,ob,	66,412	66	4,36	66,412	67	4,836
			73,333	73	8,024	73,333	74	8,524
Délka prostředníku [mm] (8)	Věk,skup,	20-29	80,301	80	5,992	80,042	80	5,808
		30-39	79,877	80	5,707	79,591	80	5,688
		40-49	79,473	79	5,626	79,387	79	5,502
		50-59	78,552	78	5,699	78,699	79	5,433
		60-69	76,179	77	6,033	76,806	77	6,311
		70-79	75,245	76	4,746	74,816	74	4,503
	BMI skup,	ideální	79,004	79	5,834	78,722	79	5,771
		mírná ob,	79,236	79	5,748	79,463	79	5,401
		nadváha	79,771	80	6,112	79,843	79	5,793
		podváha střed,ob,	76,235	77	6,582	75,529	77	6,875
			78,267	80	4,559	77,333	78	5,136

Tabulka B.4: Testy normality - délka ruky

Roz.	Skupina	Kat.	n	Pravá ruka						Levá ruka					
				JB	JB (p-h)	Lilite	Lilite (p-h)	Shar	Shar (p-h)	JB	JB (p-h)	Lilite	Lilite (p-h)	Shar	Shar (p-h)
	Typ práce	K	445	10,173	0,006	0,06	<0,001	0,986	<0,001	11,344	0,003	0,074	<0,001	0,985	<0,001
		M	529	2,652	0,266	0,044	0,015	0,996	0,166	2,934	0,231	0,035	0,113	0,996	0,159
	Pohlaví	F	432	1,863	0,394	0,056	0,002	0,995	0,161	1,706	0,426	0,045	0,036	0,994	0,115
		M	542	0,544	0,762	0,049	0,004	0,996	0,252	1,264	0,531	0,039	0,046	0,996	0,177
	Dominance	L	84	2,13	0,345	0,066	0,482	0,981	0,267	2,422	0,298	0,054	0,788	0,981	0,241
		R	890	8,953	0,011	0,049	<0,001	0,994	0,002	10,139	0,006	0,049	<0,001	0,993	<0,001
1	Věk, skup,	20-29	355	5,56	0,062	0,053	0,017	0,989	0,01	5,618	0,06	0,062	0,002	0,99	0,016
		30-39	154	1,149	0,563	0,052	0,384	0,994	0,744	1,578	0,454	0,047	0,572	0,992	0,513
		40-49	186	3,577	0,167	0,067	0,04	0,985	0,052	3,187	0,187	0,062	0,076	0,987	0,087
		50-59	163	1,716	0,424	0,067	0,068	0,992	0,446	1,874	0,392	0,076	0,023	0,989	0,207
		60-69	67	2,128	0,345	0,086	0,257	0,974	0,174	3,254	0,196	0,093	0,165	0,966	0,064
		70-79	49	0,476	0,788	0,092	0,381	0,983	0,698	0,246	0,884	0,103	0,212	0,981	0,62
	BMI skup,	ideální	474	4,992	0,082	0,056	0,001	0,994	0,053	4,732	0,094	0,062	<0,001	0,993	0,019
		mírná ob, nadváha podváha střed.ob,	123	1,62	0,445	0,065	0,228	0,989	0,413	4,25	0,119	0,076	0,076	0,978	0,043
	BMI skup,	345	5,728	0,057	0,058	0,007	0,99	0,015	5,687	0,058	0,061	0,003	0,99	0,022	
		17	0,718	0,698	0,134	0,573	0,952	0,491	0,746	0,689	0,134	0,579	0,96	0,631	
			15	0,995	0,608	0,118	0,829	0,944	0,439	0,38	0,827	0,114	0,867	0,981	0,976

Tabulka B.5: Testy normality - délka dlaně

Roz.	Skupina	Kat.	n	Pravá ruka						Levá ruka					
				JB	JB (p-h)	Lilite	Lilite (p-h)	Shar	Shar (p-h)	JB	JB (p-h)	Lilite	Lilite (p-h)	Shar	Shar (p-h)
	Typ práce	K	445	12,406	0,002	0,06	<0,001	0,982	<0,001	12,12	0,002	0,072	<0,001	0,983	<0,001
		M	529	2,094	0,351	0,054	<0,001	0,993	0,02	6,757	0,034	0,057	<0,001	0,99	<0,001
	Pohlaví	F	432	0,858	0,651	0,065	<0,001	0,993	0,034	4,22	0,121	0,074	<0,001	0,99	0,005
		M	542	0,646	0,724	0,046	0,008	0,996	0,144	1,93	0,381	0,057	<0,001	0,994	0,046
	Dominance	L	84	0,9	0,638	0,112	0,011	0,976	0,116	0,739	0,691	0,099	0,04	0,982	0,301
		R	890	11,992	0,002	0,056	<0,001	0,992	<0,001	17,191	<0,001	0,061	<0,001	0,989	<0,001
2	Věk, skup.	20-29	355	6,076	0,048	0,049	0,039	0,989	0,011	5,35	0,069	0,057	0,008	0,99	0,02
		30-39	154	0,383	0,826	0,053	0,375	0,995	0,869	0,841	0,657	0,055	0,307	0,992	0,509
		40-49	186	2,502	0,286	0,079	0,007	0,988	0,129	3,811	0,149	0,078	0,007	0,986	0,054
		50-59	163	2,617	0,27	0,08	0,012	0,988	0,16	4,942	0,085	0,075	0,026	0,979	0,013
		60-69	67	2,928	0,231	0,102	0,081	0,966	0,064	3,038	0,219	0,134	0,004	0,958	0,024
		70-79	49	0,117	0,943	0,084	0,532	0,985	0,794	2,224	0,329	0,13	0,036	0,974	0,36
	ideální mírná ob., nadváha podváha střed.ob.	474	5,161	0,076	0,052	0,004	0,992	0,011	9,3	0,01	0,071	<0,001	0,987	<0,001	
		123	1,92	0,383	0,084	0,034	0,989	0,466	3,584	0,167	0,1	0,004	0,978	0,038	
	BMI skup.	345	7,108	0,029	0,071	<0,001	0,987	0,004	6,54	0,038	0,06	0,004	0,988	0,007	
		17	1,077	0,584	0,118	0,766	0,948	0,423	2,571	0,277	0,248	0,007	0,908	0,093	
			15	0,525	0,769	0,092	0,982	0,976	0,741	0,69	0,137	0,631	0,95	0,53	

Tabulka B.6: Testy normality - šířka dlaně

Roz.	Skupina	Kat.	n	Pravá ruka						Levá ruka					
				JB	JB (p-h)	Lilie	Lilie (p-h)	Shar	Shar (p-h)	JB	JB (p-h)	Lilie	Lilie (p-h)	Shar	Shar (p-h)
	Typ práce	K	445	7,617	0,022	0,083	<0,001	0,987	<0,001	10,519	0,005	0,077	<0,001	0,985	<0,001
		M	529	9,229	0,01	0,076	<0,001	0,988	<0,001	8,721	0,013	0,06	<0,001	0,99	<0,001
	Pohlaví	F	432	3,585	0,167	0,06	<0,001	0,99	0,005	1,106	0,575	0,077	<0,001	0,991	0,007
		M	542	41,081	<0,001	0,077	<0,001	0,98	<0,001	6,484	0,039	0,068	<0,001	0,99	0,001
	Dominance	L	84	4,206	0,122	0,108	0,016	0,962	0,013	1,326	0,515	0,07	<0,001	0,988	0,611
		R	890	12,365	0,002	0,067	<0,001	0,99	<0,001	16,955	<0,001	0,067	<0,001	0,989	<0,001
3	Věk,skup,	20-29	355	8,432	0,015	0,075	<0,001	0,985	0,001	10,292	0,006	0,095	<0,001	0,98	<0,001
		30-39	154	4,338	0,114	0,099	<0,001	0,979	0,019	1,464	0,481	0,075	0,032	0,991	0,455
		40-49	186	7,881	0,019	0,125	<0,001	0,962	<0,001	4,669	0,097	0,1	<0,001	0,981	0,013
		50-59	163	5,164	0,076	0,093	0,001	0,975	0,004	7,247	0,027	0,083	0,008	0,968	<0,001
		60-69	67	0,74	0,691	0,106	0,059	0,974	0,173	2,724	0,256	0,089	0,213	0,969	0,098
		70-79	49	3,492	0,175	0,127	0,045	0,933	0,008	3,373	0,185	0,16	0,003	0,954	0,053
	BMI skup,	ideální	474	7,517	0,023	0,082	<0,001	0,988	<0,001	11,666	0,003	0,082	<0,001	0,984	<0,001
		mírná	123	4,473	0,107	0,074	0,094	0,977	0,031	3,596	0,166	0,085	0,031	0,98	0,068
		ob, nadváha podváha	345	5,828	0,054	0,08	<0,001	0,985	<0,001	5,578	0,061	0,066	<0,001	0,988	0,007
	střed,ob,	17	11,377	0,003	0,228	0,019	0,833	0,006	2,141	0,343	0,136	0,557	0,942	0,343	
		15	1,051	0,591	0,151	0,474	0,937	0,343	0,572	0,751	0,13	0,714	0,969	0,837	

Tabulka B.7: Testy normality - obvod dlaně

Roz.	Skupina	Kat.	n	Pravá ruka						Levá ruka					
				JB	JB (p-h)	Lilie	Lilie (p-h)	Shar	Shar (p-h)	JB	JB (p-h)	Lilie	Lilie (p-h)	Shar	Shar (p-h)
	Typ práce	K	445	10,265	0,006	0,099	<0,001	0,983	<0,001	9,148	0,01	0,074	<0,001	0,986	<0,001
		M	529	14,152	<0,001	0,061	<0,001	0,986	<0,001	11,456	0,003	0,056	<0,001	0,986	<0,001
	Pohlaví	F	432	6,289	0,043	0,046	0,031	0,995	0,176	0,695	0,706	0,04	0,092	0,996	0,297
		M	542	0,462	0,794	0,047	0,006	0,996	0,168	15,483	<0,001	0,063	<0,001	0,99	0,001
	Dominance	L	84	3,164	0,206	0,105	0,024	0,967	0,031	2,02	0,364	0,114	0,009	0,981	0,239
		R	890	19,839	<0,001	0,077	<0,001	0,987	<0,001	17,246	<0,001	0,065	<0,001	0,988	<0,001
4	Věk,skup,	20-29	355	9,003	0,011	0,087	<0,001	0,981	<0,001	9,239	0,01	0,104	<0,001	0,978	<0,001
		30-39	154	2,795	0,247	0,074	0,038	0,985	0,106	2,487	0,288	0,072	0,048	0,989	0,276
		40-49	186	8,068	0,018	0,103	<0,001	0,967	<0,001	6,377	0,031	0,099	<0,001	0,973	0,001
		50-59	163	6,218	0,045	0,077	0,02	0,972	0,002	6,23	0,044	0,071	0,041	0,974	0,003
		60-69	67	4,272	0,118	0,147	0,001	0,946	0,005	3,542	0,17	0,103	0,078	0,959	0,028
		70-79	49	8,475	0,014	0,147	0,01	0,94	0,015	18,815	<0,001	0,146	0,011	0,915	0,002
	BMI skup,	ideální	474	15,944	<0,001	0,092	<0,001	0,978	<0,001	14,685	<0,001	0,083	<0,001	0,977	<0,001
		mírná	123	7,693	0,021	0,116	<0,001	0,984	<0,001	5,368	0,068	0,072	0,12	0,971	0,009
		ob, nadváha	345	3,321	0,19	0,067	<0,001	0,987	0,003	3,96	0,138	0,07	<0,001	0,988	0,007
		podváha střed.ob,	17	2,466	0,291	0,229	0,019	0,896	0,057	2,068	0,356	0,202	0,063	0,907	0,088
		15	0,847	0,655	0,163	0,353	0,954	0,591	0,864	0,649	0,166	0,325	0,954	0,591	

Tabulka B.8: Testy normality - délka ukazováku

Roz.	Skupina	Kat.	n	Pravá ruka						Levá ruka					
				JB	JB (p-h)	Lilie	Lilie (p-h)	Shar	Shar (p-h)	JB	JB (p-h)	Lilie	Lilie (p-h)	Shar	Shar (p-h)
	Typ práce	K	445	2,867	0,238	0,055	0,003	0,992	0,023	2,791	0,248	0,05	0,01	0,993	0,04
		M	529	2,878	0,237	0,06	<0,001	0,992	0,009	4,984	0,083	0,062	<0,001	0,992	0,008
	Pohlaví	F	432	1,679	0,432	0,062	<0,001	0,99	0,006	3,802	0,149	0,059	0,001	0,989	0,003
		M	542	5,672	0,059	0,064	<0,001	0,99	0,001	5,091	0,078	0,064	<0,001	0,992	0,007
	Dominance	L	84	1,303	0,521	0,088	0,109	0,987	0,564	0,542	0,763	0,099	0,039	0,981	0,27
		R	890	3,624	0,163	0,055	<0,001	0,994	<0,001	2,374	0,305	0,051	<0,001	0,994	0,002
5		20-29	355	3,582	0,167	0,07	<0,001	0,99	0,021	0,55	0,76	0,06	0,004	0,995	0,292
		30-39	154	1,114	0,573	0,073	0,042	0,988	0,205	1,426	0,49	0,073	0,043	0,988	0,225
		40-49	186	3,253	0,197	0,087	0,002	0,982	0,018	5,803	0,055	0,087	0,002	0,981	0,011
		50-59	163	1,329	0,515	0,06	0,155	0,987	0,126	0,305	0,858	0,055	0,264	0,988	0,182
		60-69	67	0,183	0,913	0,07	0,566	0,987	0,697	1,223	0,542	0,085	0,268	0,977	0,238
	70-79	49	0,466	0,792	0,113	0,125	0,983	0,713	0,631	0,729	0,077	0,672	0,985	0,793	
	ideální	474	2,701	0,259	0,049	0,008	0,992	0,013	0,006	0,997	0,05	0,007	0,995	0,159	
		123	0,64	0,726	0,052	0,567	0,989	0,466	2,147	0,342	0,072	0,127	0,985	0,208	
BMI skup,	ob, nadváha podváha	345	1,943	0,378	0,063	0,002	0,992	0,045	3,089	0,213	0,058	0,007	0,991	0,026	
		17	0,619	0,734	0,155	0,342	0,95	0,46	0,248	0,883	0,148	0,414	0,957	0,574	
		15	0,323	0,851	0,173	0,267	0,95	0,522	0,325	0,85	0,141	0,579	0,955	0,604	
	střed,ob,														

Tabulka B.9: Testy normality - délka palce

Roz.	Skupina	Kat.	n	Pravá ruka						Levá ruka					
				JB	JB (p-h)	Lilie	Lilie (p-h)	Shar	Shar (p-h)	JB	JB (p-h)	Lilie	Lilie (p-h)	Shar	Shar (p-h)
	Typ práce	K	445	3,99	0,136	0,058	<0,001	0,991	0,007	1,677	0,432	0,059	<0,001	0,992	0,018
		M	529	7,082	0,029	0,066	<0,001	0,99	0,001	3,744	0,154	0,06	<0,001	0,992	0,009
	Pohlaví	F	432	7,226	0,027	0,072	<0,001	0,989	0,002	9,391	0,009	0,069	<0,001	0,988	0,002
		M	542	5,836	0,054	0,07	<0,001	0,991	0,002	2,376	0,305	0,065	<0,001	0,993	0,011
	Dominance	L	84	0,235	0,889	0,071	0,372	0,986	0,489	0,82	0,664	0,067	0,453	0,99	0,777
		R	890	6,034	0,049	0,056	<0,001	0,993	<0,001	4,776	0,092	0,054	<0,001	0,993	<0,001
6		20-29	355	2,829	0,243	0,06	0,003	0,991	0,036	3,801	0,15	0,064	0,001	0,991	0,024
		30-39	154	1,061	0,588	0,078	0,023	0,99	0,322	1,814	0,404	0,072	0,047	0,988	0,225
		40-49	186	0,863	0,65	0,077	0,01	0,987	0,085	4,41	0,11	0,09	<0,001	0,982	0,02
		50-59	163	3,626	0,163	0,082	0,009	0,981	0,027	0,59	0,745	0,071	0,043	0,992	0,453
		60-69	67	1,947	0,378	0,11	0,045	0,975	0,184	0,21	0,9	0,094	0,146	0,983	0,476
	70-79	49	0,616	0,735	0,068	0,825	0,984	0,736	2,015	0,365	0,114	0,115	0,963	0,124	
BMI skup,		ideální	474	2,356	0,308	0,07	<0,001	0,992	0,017	4,798	0,091	0,069	<0,001	0,991	0,007
		mírná	123	3,457	0,178	0,081	0,047	0,979	0,053	1,113	0,573	0,07	0,146	0,989	0,402
		ob,	345	3,005	0,223	0,058	0,008	0,988	0,008	0,262	0,877	0,062	0,003	0,993	0,112
		nadváha	17	4,353	0,113	0,165	0,252	0,924	0,173	18,764	<0,001	0,227	0,02	0,819	0,004
	střed,ob,	15	0,796	0,672	0,123	0,788	0,954	0,588	2,913	0,233	0,141	0,589	0,921	0,2	

Tabulka B.10: Testy normality - délka pro stisk spouště

Roz.	Skupina	Kat.	n	Pravá ruka				Levá ruka							
				JB	JB (p-h)	Lilie	Lilie (p-h)	Shar	Shar (p-h)	JB	JB (p-h)	Lilie	Lilie (p-h)	Shar	Shar (p-h)
	Typ práce	K	445	6,586	0,037	0,065	<0,001	0,989	0,001	7,279	0,026	0,076	<0,001	0,989	0,003
		M	529	7,861	0,02	0,053	0,001	0,989	<0,001	7,099	0,029	0,049	0,004	0,991	0,002
	Pohlaví	F	432	1,121	0,571	0,064	<0,001	0,993	0,056	4,657	0,097	0,054	0,004	0,991	0,007
		M	542	3,307	0,191	0,056	<0,001	0,993	0,01	1,226	0,542	0,046	0,009	0,995	0,082
	Dominance	L	84	1,796	0,407	0,072	0,343	0,982	0,285	3,162	0,206	0,09	0,086	0,968	0,033
		R	890	13,15	0,001	0,055	<0,001	0,99	<0,001	12,701	0,002	0,052	<0,001	0,991	<0,001
7	Věk, skup,	20-29	355	3,186	0,203	0,047	0,057	0,992	0,044	3,782	0,151	0,05	0,032	0,991	0,037
		30-39	154	1,274	0,529	0,076	0,029	0,987	0,145	2,739	0,254	0,059	0,204	0,986	0,11
		40-49	186	2,455	0,293	0,074	0,014	0,986	0,063	3,542	0,17	0,061	0,092	0,986	0,062
		50-59	163	5,303	0,071	0,065	0,087	0,979	0,014	2,887	0,236	0,056	0,245	0,985	0,07
		60-69	67	2,943	0,23	0,128	0,008	0,955	0,016	1,098	0,577	0,084	0,288	0,986	0,649
		70-79	49	9,833	0,007	0,156	0,004	0,936	0,01	3,959	0,138	0,143	0,013	0,942	0,018
	BMI skup,	ideální	474	7,372	0,025	0,062	<0,001	0,99	0,002	10,883	0,004	0,064	<0,001	0,987	<0,001
		mírná ob, nadváha podváha střed, ob,	123	5,451	0,066	0,086	0,026	0,97	0,008	1,437	0,488	0,068	0,174	0,99	0,507
	BMI skup,	nadváha	345	4,702	0,095	0,057	0,01	0,989	0,012	1,698	0,428	0,051	0,029	0,993	0,128
		podváha	17	0,526	0,769	0,137	0,544	0,978	0,938	0,822	0,663	0,172	0,197	0,953	0,506
		střed, ob,	15	0,205	0,902	0,135	0,647	0,969	0,838	1,568	0,457	0,23	0,031	0,862	0,025

Tabulka B.11: Testy normality - délka prostředníku

Roz.	Skupina	Kat.	n	Pravá ruka						Levá ruka					
				JB	JB (p-h)	Lilie	Lilie (p-h)	Shar	Shar (p-h)	JB	JB (p-h)	Lilie	Lilie (p-h)	Shar	Shar (p-h)
	Typ práce	K	445	3,754	0,153	0,047	0,02	0,994	0,056	4,258	0,119	0,054	0,004	0,993	0,049
		M	529	2,76	0,252	0,065	<0,001	0,993	0,01	0,849	0,654	0,049	0,004	0,995	0,071
	Pohlaví	F	432	1,221	0,543	0,054	0,004	0,991	0,012	0,275	0,871	0,066	<0,001	0,993	0,06
		M	542	1,133	0,568	0,054	<0,001	0,995	0,093	0,879	0,644	0,053	0,001	0,994	0,038
	Dominance	L	84	1,33	0,514	0,086	0,13	0,986	0,523	1,627	0,443	0,074	0,302	0,978	0,171
		R	890	4,035	0,133	0,05	<0,001	0,995	0,004	2,19	0,334	0,048	<0,001	0,996	0,011
8	Věk,skup,	20-29	355	1,853	0,396	0,056	0,009	0,994	0,157	2,003	0,367	0,066	<0,001	0,994	0,139
		30-39	154	1,293	0,524	0,07	0,065	0,992	0,542	2,575	0,276	0,074	0,04	0,987	0,161
		40-49	186	7,146	0,028	0,117	<0,001	0,973	0,001	1,979	0,372	0,066	0,048	0,99	0,219
		50-59	163	0,861	0,65	0,066	0,079	0,989	0,218	1,519	0,468	0,07	0,05	0,989	0,227
		60-69	67	2,187	0,335	0,09	0,193	0,965	0,056	0,491	0,782	0,062	0,764	0,988	0,743
	70-79	49	5,024	0,081	0,123	0,062	0,96	0,098	0,26	0,878	0,097	0,295	0,983	0,695	
	BMI skup,	ideální	474	2,24	0,326	0,048	0,012	0,994	0,068	1,45	0,484	0,054	0,002	0,994	0,072
		mírná ob, nadváha podváha střed,ob,	123	0,967	0,617	0,078	0,061	0,988	0,376	1,984	0,371	0,082	0,04	0,984	0,164
			345	2,772	0,25	0,068	<0,001	0,992	0,059	0,608	0,738	0,052	0,025	0,995	0,318
			17	0,463	0,793	0,124	0,691	0,971	0,844	0,537	0,765	0,134	0,57	0,96	0,631
			15	0,308	0,857	0,207	0,082	0,933	0,297	1,287	0,526	0,191	0,151	0,905	0,115

Tabulka B.12: Korelační koeficienty - rozměry s výškou

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
výška x (1) pravá	F	0,616	<0,001	0,441	<0,001	0,609	<0,001
	M	0,66	<0,001	0,479	<0,001	0,656	<0,001
výška x (2) pravá	F	0,553	<0,001	0,388	<0,001	0,542	<0,001
	M	0,591	<0,001	0,424	<0,001	0,581	<0,001
výška x (3) pravá	F	0,254	<0,001	0,166	<0,001	0,239	<0,001
	M	0,326	<0,001	0,217	<0,001	0,311	<0,001
výška x (4) pravá	F	0,281	<0,001	0,184	<0,001	0,271	<0,001
	M	0,34	<0,001	0,213	<0,001	0,307	<0,001
výška x (5) pravá	F	0,509	<0,001	0,362	<0,001	0,505	<0,001
	M	0,465	<0,001	0,337	<0,001	0,468	<0,001
výška x (6) pravá	F	0,365	<0,001	0,249	<0,001	0,353	<0,001
	M	0,279	<0,001	0,193	<0,001	0,277	<0,001
výška x (7) pravá	F	0,422	<0,001	0,298	<0,001	0,427	<0,001
	M	0,362	<0,001	0,253	<0,001	0,364	<0,001
výška x (8) pravá	F	0,473	<0,001	0,344	<0,001	0,48	<0,001
	M	0,53	<0,001	0,381	<0,001	0,532	<0,001
výška x (1) levá	F	0,615	<0,001	0,427	<0,001	0,595	<0,001
	M	0,63	<0,001	0,451	<0,001	0,624	<0,001
výška x (2) levá	F	0,542	<0,001	0,379	<0,001	0,527	<0,001
	M	0,552	<0,001	0,38	<0,001	0,53	<0,001
výška x (3) levá	F	0,296	<0,001	0,194	<0,001	0,282	<0,001
	M	0,314	<0,001	0,206	<0,001	0,298	<0,001
výška x (4) levá	F	0,28	<0,001	0,184	<0,001	0,27	<0,001
	M	0,33	<0,001	0,21	<0,001	0,303	<0,001
výška x (5) levá	F	0,508	<0,001	0,369	<0,001	0,512	<0,001
	M	0,473	<0,001	0,348	<0,001	0,487	<0,001
výška x (6) levá	F	0,32	<0,001	0,222	<0,001	0,316	<0,001
	M	0,323	<0,001	0,238	<0,001	0,337	<0,001
výška x (7) levá	F	0,421	<0,001	0,293	<0,001	0,421	<0,001
	M	0,306	<0,001	0,214	<0,001	0,311	<0,001
výška x (8) levá	F	0,475	<0,001	0,33	<0,001	0,461	<0,001
	M	0,524	<0,001	0,373	<0,001	0,528	<0,001

Tabulka B.13: Korelační koeficienty - rozměry mezi sebou

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
(1) pravá x levá	F	0,922	<0,001	0,789	<0,001	0,924	<0,001
	M	0,929	<0,001	0,785	<0,001	0,923	<0,001
(2) pravá x levá	F	0,831	<0,001	0,666	<0,001	0,823	<0,001
	M	0,873	<0,001	0,715	<0,001	0,864	<0,001
(3) pravá x levá	F	0,856	<0,001	0,719	<0,001	0,852	<0,001
	M	0,872	<0,001	0,742	<0,001	0,876	<0,001
(4) pravá x levá	F	0,906	<0,001	0,749	<0,001	0,898	<0,001
	M	0,901	<0,001	0,738	<0,001	0,891	<0,001
(5) pravá x levá	F	0,849	<0,001	0,706	<0,001	0,846	<0,001
	M	0,87	<0,001	0,736	<0,001	0,868	<0,001
(6) pravá x levá	F	0,726	<0,001	0,567	<0,001	0,717	<0,001
	M	0,795	<0,001	0,614	<0,001	0,768	<0,001
(7) pravá x levá	F	0,742	<0,001	0,571	<0,001	0,734	<0,001
	M	0,767	<0,001	0,599	<0,001	0,766	<0,001
(8) pravá x levá	F	0,797	<0,001	0,645	<0,001	0,799	<0,001
	M	0,791	<0,001	0,64	<0,001	0,803	<0,001

Tabulka B.14: Parciální korelační koeficienty po vyřazení výšky - rozměry mezi sebou

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
(1) pravá x levá	F	0,874	<0,001	0,74	<0,001	0,88	<0,001
	M	0,879	<0,001	0,727	<0,001	0,872	<0,001
(2) pravá x levá	F	0,758	<0,001	0,608	<0,001	0,752	<0,001
	M	0,813	<0,001	0,662	<0,001	0,805	<0,001
(3) pravá x levá	F	0,845	<0,001	0,71	<0,001	0,842	<0,001
	M	0,858	<0,001	0,73	<0,001	0,863	<0,001
(4) pravá x levá	F	0,898	<0,001	0,741	<0,001	0,89	<0,001
	M	0,888	<0,001	0,725	<0,001	0,88	<0,001
(5) pravá x levá	F	0,797	<0,001	0,66	<0,001	0,792	<0,001
	M	0,833	<0,001	0,701	<0,001	0,829	<0,001
(6) pravá x levá	F	0,69	<0,001	0,542	<0,001	0,682	<0,001
	M	0,776	<0,001	0,596	<0,001	0,746	<0,001
(7) pravá x levá	F	0,686	<0,001	0,529	<0,001	0,675	<0,001
	M	0,739	<0,001	0,576	<0,001	0,738	<0,001
(8) pravá x levá	F	0,738	<0,001	0,6	<0,001	0,742	<0,001
	M	0,711	<0,001	0,58	<0,001	0,726	<0,001

Tabulka B.15: Korelační koeficienty - porovnání s věkem

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
věk x (3) pravá	F	0,317	<0,001	0,214	<0,001	0,303	<0,001
	M	0,25	<0,001	0,195	<0,001	0,277	<0,001
věk x (3) levá	F	0,25	<0,001	0,169	<0,001	0,245	<0,001
	M	0,248	<0,001	0,184	<0,001	0,262	<0,001
věk x (4) pravá	F	0,259	<0,001	0,182	<0,001	0,265	<0,001
	M	0,236	<0,001	0,181	<0,001	0,265	<0,001
věk x (4) levá	F	0,267	<0,001	0,183	<0,001	0,267	<0,001
	M	0,22	<0,001	0,165	<0,001	0,238	<0,001
věk x (6) pravá	F	0,205	<0,001	0,111	<0,001	0,158	<0,001
	M	0,191	<0,001	0,144	<0,001	0,209	<0,001
věk x (6) levá	F	0,188	<0,001	0,119	<0,001	0,171	<0,001
	M	0,131	0,002	0,091	0,002	0,132	0,002

Tabulka B.16: Parciální korelační koeficienty - porovnání s věkem

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
věk x (3) pravá	F	0,425	<0,001	0,257	<0,001	0,403	<0,001
	M	0,363	<0,001	0,238	<0,001	0,375	<0,001
věk x (3) levá	F	0,37	<0,001	0,218	<0,001	0,359	<0,001
	M	0,357	<0,001	0,223	<0,001	0,354	<0,001
věk x (4) pravá	F	0,373	<0,001	0,229	<0,001	0,376	<0,001
	M	0,354	<0,001	0,222	<0,001	0,361	<0,001
věk x (4) levá	F	0,381	<0,001	0,23	<0,001	0,379	<0,001
	M	0,332	<0,001	0,205	<0,001	0,329	<0,001
věk x (6) pravá	F	0,352	<0,001	0,172	<0,001	0,295	<0,001
	M	0,282	<0,001	0,18	<0,001	0,289	<0,001
věk x (6) levá	F	0,312	<0,001	0,173	<0,001	0,293	<0,001
	M	0,232	<0,001	0,134	<0,001	0,226	<0,001

Tabulka B.17: Korelační koeficienty - porovnání se šířkou

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
(3) x (1) pravá	F	0,517	<0,001	0,367	<0,001	0,505	<0,001
	M	0,572	<0,001	0,403	<0,001	0,552	<0,001
(3) x (1) levá	F	0,533	<0,001	0,386	<0,001	0,526	<0,001
	M	0,579	<0,001	0,399	<0,001	0,548	<0,001
(3) x (2) pravá	F	0,488	<0,001	0,349	<0,001	0,477	<0,001
	M	0,552	<0,001	0,401	<0,001	0,549	<0,001
(3) x (2) levá	F	0,509	<0,001	0,383	<0,001	0,516	<0,001
	M	0,559	<0,001	0,391	<0,001	0,532	<0,001
(3) x (4) pravá	F	0,844	<0,001	0,68	<0,001	0,834	<0,001
	M	0,828	<0,001	0,679	<0,001	0,834	<0,001
(3) x (4) levá	F	0,86	<0,001	0,702	<0,001	0,851	<0,001
	M	0,841	<0,001	0,681	<0,001	0,84	<0,001

Tabulka B.18: Parciální korelační koeficienty - porovnání se šířkou

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
(3) x (1) pravá	F	0,473	<0,001	0,332	<0,001	0,467	<0,001
	M	0,503	<0,001	0,349	<0,001	0,486	<0,001
(3) x (1) levá	F	0,466	<0,001	0,342	<0,001	0,464	<0,001
	M	0,517	<0,001	0,35	<0,001	0,485	<0,001
(3) x (2) pravá	F	0,431	<0,001	0,313	<0,001	0,425	<0,001
	M	0,471	<0,001	0,349	<0,001	0,476	<0,001
(3) x (2) levá	F	0,434	<0,001	0,341	<0,001	0,45	<0,001
	M	0,488	<0,001	0,345	<0,001	0,462	<0,001
(3) x (4) pravá	F	0,832	<0,001	0,67	<0,001	0,823	<0,001
	M	0,807	<0,001	0,663	<0,001	0,817	<0,001
(3) x (4) levá	F	0,847	<0,001	0,691	<0,001	0,839	<0,001
	M	0,823	<0,001	0,667	<0,001	0,824	<0,001

Tabulka B.19: Korelační koeficienty - porovnání s BMI

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
BMI x (3) pravá	F	0,334	<0,001	0,241	<0,001	0,346	<0,001
	M	0,425	<0,001	0,294	<0,001	0,42	<0,001
BMI x (3) levá	F	0,338	<0,001	0,235	<0,001	0,336	<0,001
	M	0,44	<0,001	0,296	<0,001	0,424	<0,001
BMI x (4) pravá	F	0,449	<0,001	0,327	<0,001	0,469	<0,001
	M	0,479	<0,001	0,325	<0,001	0,467	<0,001
BMI x (4) levá	F	0,473	<0,001	0,333	<0,001	0,477	<0,001
	M	0,471	<0,001	0,321	<0,001	0,463	<0,001

Tabulka B.20: Parciální korelační koeficienty - porovnání s BMI

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
BMI x (3) pravá	F	0,4	<0,001	0,268	<0,001	0,413	<0,001
	M	0,485	<0,001	0,316	<0,001	0,475	<0,001
BMI x (3) levá	F	0,418	<0,001	0,268	<0,001	0,417	<0,001
	M	0,498	<0,001	0,317	<0,001	0,475	<0,001
BMI x (4) pravá	F	0,531	<0,001	0,36	<0,001	0,554	<0,001
	M	0,546	<0,001	0,348	<0,001	0,523	<0,001
BMI x (4) levá	F	0,555	<0,001	0,367	<0,001	0,562	<0,001
	M	0,535	<0,001	0,343	<0,001	0,518	<0,001

Tabulka B.21: Korelační koeficienty - porovnání se šířkou (typ práce)

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
(3) x (1) pravá	K	0,779	<0,001	0,599	<0,001	0,784	<0,001
	M	0,746	<0,001	0,577	<0,001	0,762	<0,001
(3) x (1) levá	K	0,801	<0,001	0,615	<0,001	0,797	<0,001
	M	0,746	<0,001	0,575	<0,001	0,76	<0,001
(3) x (2) pravá	K	0,773	<0,001	0,596	<0,001	0,779	<0,001
	M	0,716	<0,001	0,557	<0,001	0,739	<0,001
(3) x (2) levá	K	0,784	<0,001	0,611	<0,001	0,791	<0,001
	M	0,738	<0,001	0,576	<0,001	0,757	<0,001
(3) x (4) pravá	K	0,903	<0,001	0,751	<0,001	0,902	<0,001
	M	0,93	<0,001	0,811	<0,001	0,935	<0,001
(3) x (4) levá	K	0,911	<0,001	0,767	<0,001	0,91	<0,001
	M	0,945	<0,001	0,824	<0,001	0,948	<0,001

Tabulka B.22: Korelační koeficienty - porovnání se šířkou (typ práce)

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
(3) x (1) pravá	K	0,577	<0,001	0,439	<0,001	0,564	<0,001
	M	0,421	<0,001	0,367	<0,001	0,42	<0,001
(3) x (1) levá	K	0,592	<0,001	0,447	<0,001	0,555	<0,001
	M	0,418	<0,001	0,369	<0,001	0,422	<0,001
(3) x (2) pravá	K	0,569	<0,001	0,446	<0,001	0,56	<0,001
	M	0,378	<0,001	0,349	<0,001	0,389	<0,001
(3) x (2) levá	K	0,571	<0,001	0,46	<0,001	0,564	<0,001
	M	0,413	<0,001	0,379	<0,001	0,433	<0,001
(3) x (4) pravá	K	0,83	<0,001	0,671	<0,001	0,82	<0,001
	M	0,862	<0,001	0,739	<0,001	0,863	<0,001
(3) x (4) levá	K	0,833	<0,001	0,686	<0,001	0,821	<0,001
	M	0,89	<0,001	0,755	<0,001	0,889	<0,001

Tabulka B.23: Korelační koeficienty - porovnání s BMI (typ práce)

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
BMI x (3) pravá	K	0,408	<0,001	0,294	<0,001	0,421	<0,001
	M	0,362	<0,001	0,252	<0,001	0,361	<0,001
BMI x (3) levá	K	0,392	<0,001	0,279	<0,001	0,4	<0,001
	M	0,377	<0,001	0,261	<0,001	0,374	<0,001
BMI x (4) pravá	K	0,415	<0,001	0,304	<0,001	0,433	<0,001
	M	0,425	<0,001	0,297	<0,001	0,428	<0,001
BMI x (4) levá	K	0,407	<0,001	0,29	<0,001	0,417	<0,001
	M	0,439	<0,001	0,302	<0,001	0,435	<0,001

Tabulka B.24: Parciální korelační koeficienty - porovnání s BMI (typ práce)

Rozměr	Skup.	Pearson	Pearson p-h	Kendall	Kendall p-h	Spearman	Spearman p-h
BMI x (3) pravá	K	0,506	<0,001	0,311	<0,001	0,507	<0,001
	M	0,434	<0,001	0,263	<0,001	0,437	<0,001
BMI x (3) levá	K	0,499	<0,001	0,296	<0,001	0,494	<0,001
	M	0,462	<0,001	0,273	<0,001	0,458	<0,001
BMI x (4) pravá	K	0,569	<0,001	0,332	<0,001	0,572	<0,001
	M	0,522	<0,001	0,313	<0,001	0,526	<0,001
BMI x (4) levá	K	0,55	<0,001	0,315	<0,001	0,548	<0,001
	M	0,543	<0,001	0,319	<0,001	0,538	<0,001

Tabulka B.25: ANOVA - pohlaví

Rozměr	F test	F test p-h	Welch test	Welch test p-h	Wilcox test	Wilcox test p-h
(1) pravá	0,711	<0,001	-28,626	<0,001	22997	<0,001
(2) pravá	0,651	<0,001	-27,36	<0,001	25651	<0,001
(3) pravá	0,559	<0,001	-31,256	<0,001	19086	<0,001
(4) pravá	0,635	<0,001	-38,383	<0,001	10213,5	<0,001
(5) pravá	0,814	0,026	-16,81	<0,001	51336,5	<0,001
(6) pravá	0,676	<0,001	-14,117	<0,001	60150,5	<0,001
(7) pravá	0,595	<0,001	-17,611	<0,001	50937	<0,001
(8) pravá	0,866	0,117	-20,579	<0,001	41022	<0,001
(1) levá	0,705	<0,001	-28,703	<0,001	23136	<0,001
(2) levá	0,663	<0,001	-28,135	<0,001	24697	<0,001
(3) levá	0,605	<0,001	-33,095	<0,001	16683	<0,001
(4) levá	0,635	<0,001	-38,043	<0,001	10165,5	<0,001
(5) levá	0,703	<0,001	-17,586	<0,001	48638	<0,001
(6) levá	0,793	0,012	-14,414	<0,001	59081	<0,001
(7) levá	0,681	<0,001	-20,267	<0,001	42632,5	<0,001
(8) levá	0,859	0,098	-19,909	<0,001	42796,5	<0,001

Tabulka B.26: ANOVA - dominantní ruka

Rozměr	F test	F test p-h	Welch test	Welch test p-h	Wilcox test	Wilcox test p-h
(1) pravá	0,917	0,628	2,03	0,045	42491	0,038
(2) pravá	0,827	0,274	1,629	0,106	41481	0,096
(3) pravá	0,996	0,983	0,177	0,86	37813	0,861
(4) pravá	1,064	0,666	0,329	0,743	38227,5	0,731
(5) pravá	0,999	0,97	0,269	0,788	38443,5	0,666
(6) pravá	1,222	0,189	0,574	0,567	38824	0,557
(7) pravá	0,918	0,636	2,662	0,009	43788	0,009
(8) pravá	1,056	0,701	2,042	0,044	42683,5	0,031
(1) levá	0,916	0,624	1,13	0,261	40376,5	0,224
(2) levá	0,851	0,355	0,637	0,526	39253	0,447
(3) levá	1,059	0,689	2,387	0,019	43199,5	0,018
(4) levá	1,205	0,221	2,267	0,026	43473	0,013
(5) levá	1,062	0,677	0,889	0,376	39610,5	0,365
(6) levá	1,361	0,043	0,839	0,404	39680,5	0,35
(7) levá	1,008	0,923	-0,174	0,862	37132	0,92
(8) levá	0,912	0,607	1,507	0,135	41193	0,121

Tabulka B.27: ANOVA - typ práce

Rozměr	F test	F test p-h	Welch test	Welch test p-h	Wilcox test	Wilcox test p-h
(1) pravá	1,466	<0,001	-1,452	0,147	110720	0,11
(2) pravá	1,385	<0,001	-0,978	0,328	111993	0,191
(3) pravá	1,154	0,115	-3,839	<0,001	100873	<0,001
(4) pravá	1,161	0,1	-3,101	0,002	103946,5	0,002
(5) pravá	1,283	0,006	-1,655	0,098	110987	0,124
(6) pravá	1,224	0,026	-4,79	<0,001	96720	<0,001
(7) pravá	1,087	0,36	-1,306	0,192	111936	0,187
(8) pravá	1,312	0,003	-1,732	0,084	110463,5	0,097
(1) levá	1,355	<0,001	-1,844	0,065	109530	0,062
(2) levá	1,262	0,01	-1,348	0,178	111516	0,157
(3) levá	1,138	0,155	-4,685	<0,001	97334	<0,001
(4) levá	1,139	0,152	-3,444	<0,001	103265,5	<0,001
(5) levá	1,227	0,024	-1,895	0,058	110300	0,09
(6) levá	1,002	0,977	-2,639	0,008	106916,5	0,013
(7) levá	1,047	0,61	-1,188	0,235	112423	0,227
(8) levá	1,392	<0,001	-2,047	0,041	109345	0,056

Tabulka B.28: ANOVA - věková skupina

Rozměr	Levenův test	Levenův test p-h	ANOVA	ANOVA p-h	Kruskal-Wallis test	Kruskal-Wallis p-h
(1) pravá	2,073	0,067	10,855	<0,001	53,336	<0,001
(2) pravá	1,586	0,161	8,402	<0,001	41,033	<0,001
(3) pravá	4,154	<0,001	3,855	0,002	17,8	0,003
(4) pravá	3,62	0,003	3,832	0,002	19,423	0,002
(5) pravá	1,211	0,302	4,1	0,001	20,225	0,001
(6) pravá	2,655	0,022	2,714	0,019	12,645	0,027
(7) pravá	3,835	0,002	7,734	<0,001	42,078	<0,001
(8) pravá	2,123	0,061	10,31	<0,001	53,401	<0,001
(1) levá	2,559	0,026	9,742	<0,001	49,907	<0,001
(2) levá	1,527	0,179	5,645	<0,001	28,195	<0,001
(3) levá	3,534	0,004	4,074	0,001	20,162	0,001
(4) levá	3,58	0,003	3,58	0,003	19,102	0,002
(5) levá	1,504	0,186	4,832	<0,001	24,614	<0,001
(6) levá	2,07	0,067	0,802	0,549	3,248	0,662
(7) levá	1,514	0,183	7,443	<0,001	35,922	<0,001
(8) levá	1,007	0,412	11,747	<0,001	55,936	<0,001

Tabulka B.29: ANOVA - BMI kategorie

Rozměr	Levenův test	Levenův test p-h	ANOVA	ANOVA p-h	Kruskal-Wallis test	Kruskal-Wallis p-h
(1) pravá	0,963	0,427	8,194	<0,001	28,038	<0,001
(2) pravá	1,834	0,12	10,257	<0,001	37,202	<0,001
(3) pravá	1,694	0,149	36,654	<0,001	122,859	<0,001
(4) pravá	1,458	0,213	42,805	<0,001	141,889	<0,001
(5) pravá	1,295	0,27	2,392	0,049	7,508	0,111
(6) pravá	0,652	0,625	5,475	<0,001	22,41	<0,001
(7) pravá	1,357	0,247	6,595	<0,001	23,771	<0,001
(8) pravá	0,806	0,521	4,069	0,003	14,378	0,006
(1) levá	1,14	0,336	8,224	<0,001	28,944	<0,001
(2) levá	1,952	0,1	14,255	<0,001	53,459	<0,001
(3) levá	0,873	0,479	36,71	<0,001	119,708	<0,001
(4) levá	1,104	0,354	45,53	<0,001	145,928	<0,001
(5) levá	0,659	0,621	3,027	0,017	9,633	0,047
(6) levá	2,758	0,027	4,264	0,002	19,549	<0,001
(7) levá	2,166	0,071	7,594	<0,001	30,535	<0,001
(8) levá	0,917	0,453	2,078	0,082	6,9	0,141

Tabulka B.30: Přehled vybraných rozměrů ruky pro různé populace - ženy

Stát	Rok	n	Délka ruky (1)		Délka dlaně (2)		Šířka dlaně (3)		Délka prostředníku (8)	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
Anglie	92	1980							77,8	5,3
Brazílie	206	2017	177,3	8,91	102,7	5,21	78,17	3,83	74,6	
Čína	100	1984							77,49	3,7
ČR	433	2015	176,02	7,91	100,4	4,765	78,795	4,04	75,62	4,74
Filipíny	962	2007	179,5	3,44			92,3	6,97		
Indie	95	2003	169,6	9,4	93,6		68	5,1	76	5,7
Japonsko	768	1967							73	4
Jordánsko	120	2008	171,27	7,44	96,12		77,82	3,92	75,15	3,62
Kolumbie	126	2017	167,3	8	93,9		78,8	3,4	73,4	4,3
Mexiko	25	2005	171,8	10	98,9		77	4	72,9	5,1
Nigérie	37	2000	175,05	11,07	100,85		75,75	5,17	74,2	5,49
Thajsko	150	2004	177,22	7,04	100,45		70,16	3,29	76,77	3,24
Turecko	73	2012	172,16	8,14	97,75		69,91	3,23	74,41	3,91
USA	211	1971							78	5,1
USA (Bangladéšané)	51	2009	167,64	11,49	94,69		74,12	4,05	72,95	3,77
USA (Vietnamci)	30	1993	165	9	92,7		71	4,3	72,3	4,6

Tabulka B.31: Přehled vybraných rozměrů ruky pro různé populace - muži

Stát	Rok	n	Délka ruky (1)		Délka dlaně (2)		Šířka dlaně (3)		Délka prostředníku (8)	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
Brazílie	2123	2017	190,4	10,35	112,1	6,38	87,7	4,27	78,3	
ČR	554	2015	191,99	9,425	109,9	5,89	88,39	5,365	82,09	5,125
Filipíny	843	2007	197,5	7,82			98	4,07		
Indie (Harijána)	878	2011	185,77	6,32	105,59	4,57	84,85	2,82	79,05	4,31
Írán	529	2016	193,01	15	110,43	6,24	87,32	5,39	79	5,66
Jordánsko	115	2008	191,2	10,2	109,94		87,7	4,82	81,26	7,14
Kolumbie	129	2017	190,34	13,6	104,55	8,57	82,65	9,42	79,33	6,11
Mexiko	25	2005	185,5	7,1	108,5		85,3	4,9	78,5	4,4
Nigérie	35	2015	195,94	10,94	108,23		88,29	6,01	87,71	5,5
Turecko	92	2012	190,4	9,69	108,54		78,44	4,52	81,86	5,15
USA (Bangladéšané)	50	2009	173,15	8,38	97,34		79,39	7,16	75,81	3,39
USA (Vietnamci)	41	1993	177	12	98,8		79,2	6,9	78,2	4,5