

# PŘEHLED KURZU SEZNÁMENÍ SE SPSS

PSY252 – Statistická  
analýza dat II  
0. setkání

# DNEŠNÍ PROGRAM

Představení kurzu

Představení zdrojů

- Učebnice
- Data

Představení SPSS a základních analýz

# CÍLE KURZU

Získat praktickou schopnost provádět statistické analýzy s více než 2 proměnnými

Rozumět prezentovaným výsledkům

Korektně komunikovat výsledky analýz

# PŘEHLED TÉMAT

Seznámení se SPSS

Opáčko – analýzy z minulého semestru

Lineární regrese

Logistická regrese

Analýza rozptylu

Analýza rozptylu pro opakovaná měření

Faktorová analýza

# POŽADAVKY A ZKOUŠKA

Zpracování všech průběžných úkolů (dvojice)

- Na každý seminář
- Nebodováno ... ale možnost zisku až 5 bonusových bodů

Písemná část zkoušky

- Během hodiny a půl vypracovat zprávu z analýzy na počítači
- Se všemi zdroji, vč. googlení

Ústní část zkoušky

- Krátké setkání bez zdrojů

# UČEBNICE



- Field, A.: ***Discovering statistics using SPSS***, 4<sup>th</sup> Ed. Sage, 2013.

- <http://www.statisticshell.com/>

- <http://www.uk.sagepub.com/field4e/study/default.htm>

- Morgan, S. E., Reichert, T., Harrison, T. R.: ***From numbers to words. Reporting statistical results for the social sciences***. Allyn & Bacon, 2002.

American Psychological Association. (2001). ***Publication manual of the American Psychological Association (6th ed.)***. Washington, DC: Author.

# OPÁČKO S FIELDEM

Nelze všechno přečíst hned. Prolistujte a vracejte se podle potřeby.

Kap 1 – popisná statistika

Kap 2 – statistická indukce

Kap 3 – seznámení se SPSS

Kap 4 – Vyrábíme grafy

Kap 5 – Kontrolujeme předpoklady testů

Kap 6 – Neparametrické testy

Kap 7 – Korelace

Kap 9 – t-testy

# DATA

**Long2** – prof. Macek, adolescenti, 2 kohorty

**EU Kids Online** – celoevropský průzkum

a vlastní data...



# PŘÍKLADY ZDROJŮ DAT

(g: „free data sets“)

<http://lib.stat.cmu.edu/cgi-bin/dasl.cgi?query=Psychology&submit=Search!&metaname=topics&sort=swishrank> (data v textové podobě v odkazu za **Datafile Name:**)

[http://www.amstat.org/publications/jse/jse\\_data\\_archive.htm](http://www.amstat.org/publications/jse/jse_data_archive.htm)

StatSci.org – metazdroj -  
<http://www.statsci.org/datasets.html>

<http://www.kdnuggets.com/datasets/index.html> - metazdroj

<http://www.cessda.org/> (Některá data na objednávku, jiná přímo ke stažení)

<https://openfmri.org/data-sets> - Volná data z funkční magnetické rezonance

# SPSS

Nainstalovat z [inet.muni.cz](http://inet.muni.cz)

Import a export dat

Datová matice a orientace v datech

Data, output, syntax

Transformace proměnných

Popisné statistiky

Grafy

# ÚKOL: NASTARTOVAT

- Nainstalovat si SPSS a sehnat si Fielda a Morganovou
- Zorientovat se v EUKO a Long2 datech a jejich popisných materiálech, popř. najít si svůj datový soubor.
- Zopakovat si obsah PSY117
- Najít si partáka

# JSOU DÍVKY VÍCE MONITOROVÁNY SVÝMI RODIČI NEŽ KLUCI?

## Škála monitorování

vytvořit + vnitřní konzistence

popsat rozložení i graficky

## Pohlaví

## Test hypotézy

popis rozložení v porovnávaných skupinách

t-test nebo nějaký neparametrický test? ... předpoklady?

velikost účinku

## Formulace závěru

V souladu s naším očekáváním udávají kluci ( $M=2,72$ ,  $SD=0,56$ ) přibližně o čtvrt směrodatné odchylky nižší míru monitorování než dívky ( $M=2,86$ ;  $SD=0,58$ ),  $t(759)=3,30$ ;  $p<0,01$ ; Cohen  $d=0,24$ ; 95% CI (0,06;0,22).

# VYPOVÍDAJÍ RESPONDENTI O VŘELOSTI OBOU RODIČŮ STEJNĚ?

Škála vřelosti matky a otce

vytvořit + vnitřní konzistence

popsat rozložení i graficky

vztah mezi vřelostmi

## Test hypotézy

t-test nebo nějaký neparametrický test? ... předpoklady?

velikost účinku

## Formulace závěru

Respondenti v průměru udávají vřelost matky ( $M=3,23$ ,  $SD=0,47$ ) přibližně o čtvrt směrodatné odchylky vyšší než vřelost otce ( $M=3,08$ ;  $SD=0,58$ ),  $t(743)=7,61$ ;  $p<0,01$ ; Cohen  $d=0,29$ ; 95% CI (0,11;0,19).

# LIŠÍ SE POČET DĚTÍ V RODINÁCH RŮZNÝCH NÁRODNOSTÍ?

Národnost a Počet dětí  
četnosti a překódování

Test hypotézy

kontingční tabulka

t-test nebo nějaký neparametrický test? ... předpoklady?

velikost účinku

Formulace závěru

Rozložení kategorizovaného počtu dětí v českých rodinách se liší od počtu dětí v rodinách jiných národností ( $\chi^2(3, N=768)=9,44$ ;  $p=0,02$ ; Cramér  $V=0,11$ ). Adjustovaná rezidua indikují, že rozdíl spočívá především v menší relativní četnosti českých rodin se 3 dětmi a větší rel. četnosti českých rodin s jedináčky.

# DOBRÁ PRAXE PŘI PRÁCI S DATY

## Příprava dat

R/O čistá datová matice v XLS nebo CSV

Import do SAV a nastavit R/O (lze zaznamenat syntax)

Výpočet odvozených proměnných uložit do syntaxu

Veškeré přepočty děláme do nových proměnných

## Práce s daty

Otevřu SAV s čistými daty a spustím syntax, který mi vytvoří odvozené proměnné a uložím jako nový soubor

Dělám analýzy

Když si data nějak poškodím, vrátím se k R/O čistým datům

Když najdu v datech chybu, opravím v XLS a znovu provedu import

# PREZENTACE STATISTICKÝCH ANALÝZ (SEKCE RESULTS)

Představení dat (metoda, výpočet škál, vnitřní konzistence)

Popisné statistiky (popř. zobrazení rozložení, tabulky/grafy dle APA)

Formulace hypotézy (často pouze implicitní, žádné formality)

Zdůvodnění volby testu, popř. analytického postupu

Rekapitulace splnění předpokladů zvoleného testu

Standardní prezentace testových statistik (u jednodušších testů v textu, u složitějších modelů v tabulkách) vč. velikosti účinku (ideálně i intervalu spolehlivosti)

Interpretace výsledků testu (modelu) vzhledem k hypotéze



# 1. SEMINÁRNÍ ÚKOL

Ve dvojicích

Realizovat 3 analýzy – na t-test, na korelaci, na chíkvadrát

Z analýz sepsat zprávu v souladu s konvencemi

Odevzdat do neděle - půlnoci

DĚKUJEME ZA  
POZORNOST