

Míra sebevražd ve světě

Tomáš Jelínek

Červenec 2023

Abstract

Míra sebevražd je dobrým indikátorem duševního zdraví populace, jelikož bývají vyvrcholením závažných psychických¹⁻⁴ problémů jedince a navíc se i dobře měří. Nalezení jevů asociovaných se sebevraždami a zkoumání efektivity nástrojů s cílem snížit jejich míru je esenciální proto, abychom jim co nejvíce předcházeli. V této statistické práci se zaměřím jak na analýzu jednoho z rizik, tak na impakt duševní zdravotní péče.

Data jsou brána z volně dostupného datasetu na kaggle [Mental Health and Suicide Rates](#). Ten vychází z dat WHO - moje analýzy jsou dělány na údajích z roku 2016. Zdrojové kódy jsou volně dostupné na mém [GitHubu](#).

1 Sebevraždy u mužů a žen

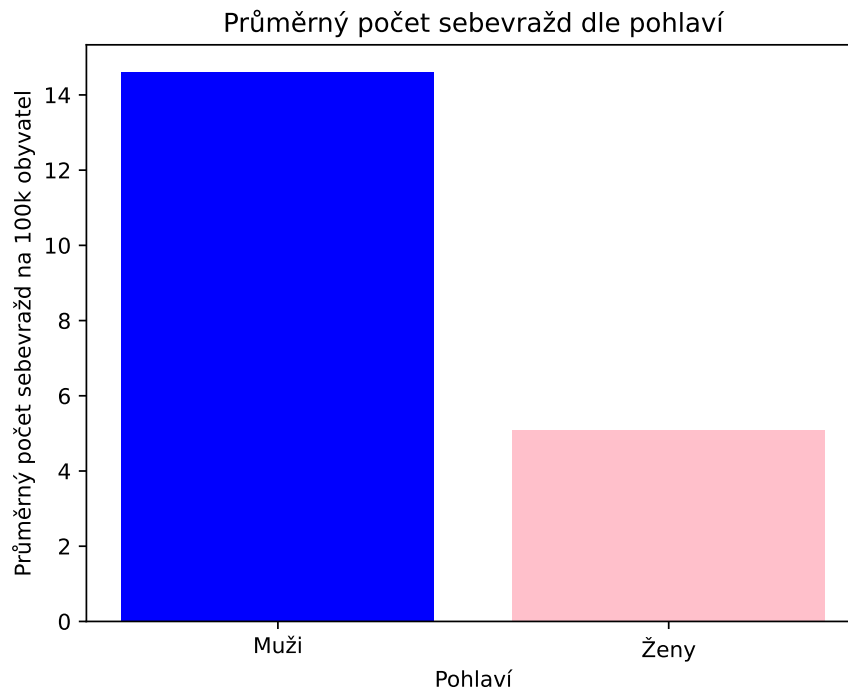
Podívejme se nyní, jestli jsou obě dvě skupiny obyvatel pod stejným rizikem. Bud' H_0 , že ženy a muži jsou ve světě pod stejným rizikem a bud' H_1 že jedno pohlaví má větší šanci na sebevraždu. Hypotézu ověřím dvouvýběrovým t-testem a jako hladinu významnosti volím $\alpha = 0.05$. T-test má určité předpoklady.

1. Nezávislost jevů - lze předpokládat, že drtivá většina sebevražd bude nezávislá na jiných.
2. Normální rozdělení - lze trochu očekávat rozdělení docela podobné normálnímu, ovšem ne přímo normální. (Extrémy a míra sebevražd nemůže být záporná nám v tom brání). Ovšem dataset je velký ($n > 30$) a tedy díky centrální limitní větě si můžeme dovolit udělat t-test.
3. Poslední podmínkou k užití je to, že variance budou hodně podobné - vzhledem k tomu, že netuším, zdalipak budou podobné, tak zvolím Welchův t-test, který tohle nepředpokládá a umožňuje je mít rozdílné.

Zde je analýza oněch dat v pythonu.

```
import numpy as np
import pandas as pd
import scipy.stats as stats
def welch_ttest():
    df = pd.read_csv("data/Age-standardized suicide rates.csv")
    df_male = df.loc[df["Sex"].str.contains("Male")]
    df_female = df.loc[df["Sex"].str.contains("Female")]
    np_males = df_male["2016"].to_numpy()
    np_females = df_female["2016"].to_numpy()
    t_statistic, p_value = stats.ttest_ind(np_males, np_females,
                                          equal_var = False)
    return(t_statistic, p_value)
```

Jako výstup dostáváme: $t_statistic = 13.4164773450928$ a $p_value = 2.8621990375255226^{-31}$. Vzhledem k tomu, že p_value je výrazně nižší než naše α , tak H_0 zamítáme. Pro zajímavost zde je barplot průměrné míry sebevražd pro obě pohlaví.



A tabulka zemí, kde byla míra sebevražd žen v roce 2016 vyšší. Ve zbylých 176 zemích převažovala u mužů.

Country	Female	Male
Antigua and Barbuda	0.9	0.0
Bangladesh	6.7	5.5
China	8.3	7.9
Lesotho	32.6	22.7
Morocco	3.6	2.5
Myanmar	9.8	6.3
Pakistan	3.1	3.0

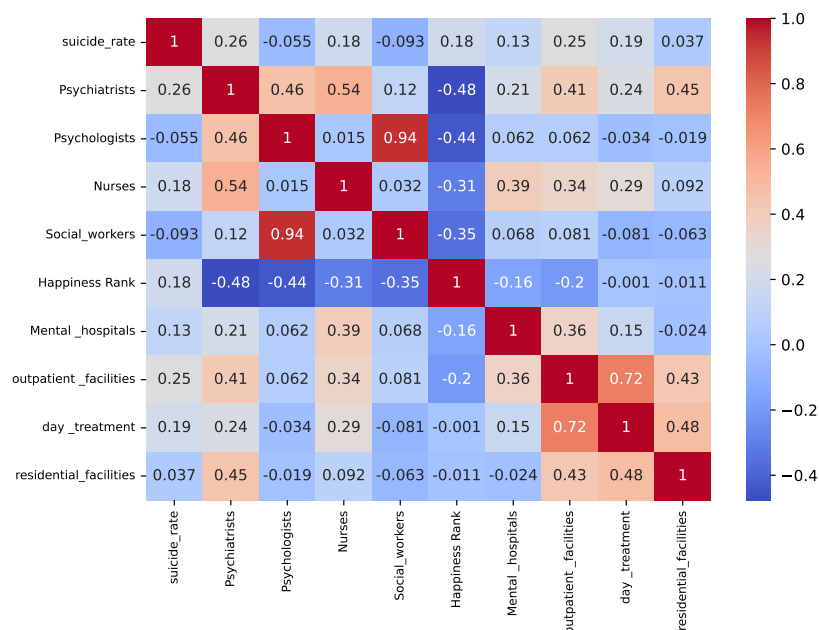
2 Léčebné instituce, štěstí a míra sebevražd

2.1 Léčebné instituce

Mezi nástroje pro zlepšování mentálního zdraví patří odoborná péče od psychiatrů, psychologů, sociálních pracovníků atd. Bude tedy míra sebevražd lepší v zemích, kde je péče dostupnější? Vyrobneme si korelační matici:

```
#Imports like in previous code snippet
import seaborn as sns
df = load_merged_data() #more about this on my GitHub
df = df.loc[df["Sex"].str.contains("Both sexes")]
df = df.rename(columns={"2016": "suicide_rate"})
df = df.loc[:, ["suicide_rate", "Psychiatrists", "Psychologists",
               "Nurses", "Social_workers",
               "Happiness Rank", "Mental
               _hospitals", "outpatient
               _facilities", "day
               _treatment", "
               residential_facilities"]]

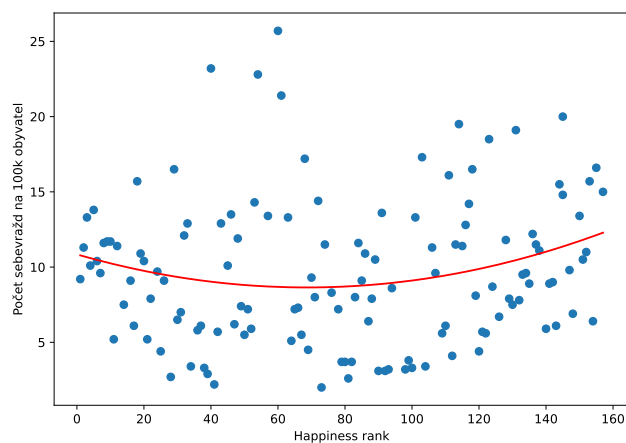
df.to_csv("correlation_matrix.csv")
corr_matrix = df.corr()
fig, ax = plt.subplots(figsize=(8, 6))
sns.heatmap(corr_matrix, annot=True, cmap="coolwarm", ax=ax)
plt.show()
```



Počty psychiatrů/psychologů, mentálních institucí atd. na hlavu nemají sami o sobě téměř žádnou korelaci s mírou sebevražd, i když by člověk na první pohled asi očekával, že budou mít negativní.

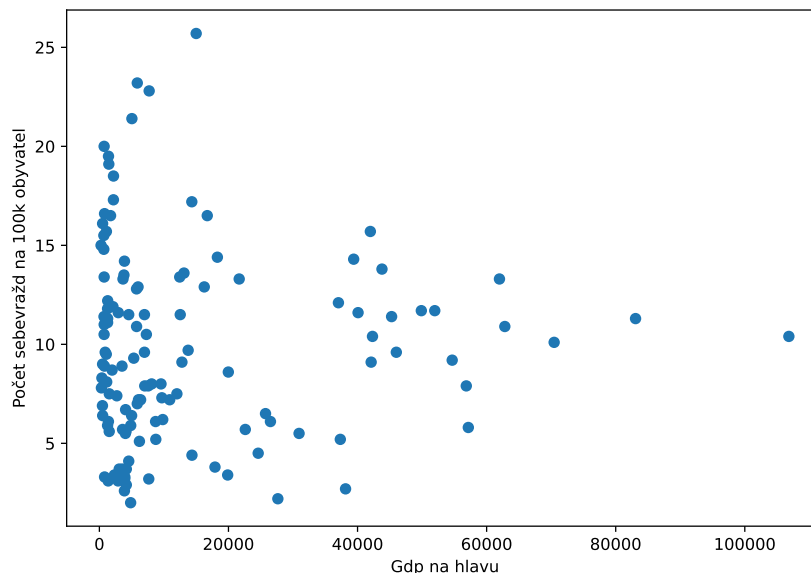
2.2 Štěstí

Zajímavé je, že ani rank státu v žebříčku štěstí obyvatel s tím nekoreluje, byť je v datech patrná mírná polynomiální závislost. (Data štěstí jsou z tohoto [datasetu](#)).



2.3 GDP

Člověk někdy na [internetu narazí na tvrzení](#), že v bohatších zemích páchají lidé častěji sebevraždy.



Z plotu dat je vidět, že u tohoto tvrzení musí být člověk opatrný. Byť je v bohatších zemích míra sebevražd značně větší než v některých chudších. Tak ale je opět mnoho chudých zemí, kde je míra sebevražd vyšší nebo stejná jak v těch bohatých.

3 Slovo závěrem

Pokud vás případně zajímají další statistiky, jako je například vývoj sebevražd v čase, jiné regrese, či dělat věci dle regionů (Evropa, Afrika...). Můžete si je nechat udělat na mém [GitHubu](#).

Reference

- ¹ J. M. Bertolote and A. Fleischmann. A global perspective in the epidemiology of suicide. *Suicidologi*, 7(2), 2002.
- ² J. M. Bertolote, A. Fleischmann, D. De Leo, and D. Wasserman. Psychiatric diagnoses and suicide: revisiting the evidence. *Crisis*, 25(4):147–155, 2004.
- ³ B. Chang, D. Gitlin, and R. Patel. The depressed patient and suicidal patient in the emergency department: evidence-based management and treatment strategies. *Emergency medicine practice*, 13(9):1–23, 2011.

⁴ A. J. Ferrari, R. E. Norman, G. Freedman, A. J. Baxter, J. E. Pirkis, M. G. Harris, A. Page, E. Carnahan, L. Degenhardt, T. Vos, et al. The burden attributable to mental and substance use disorders as risk factors for suicide: findings from the global burden of disease study 2010. *PloS one*, 9(4):e91936, 2014.