

Vplyv sociálnych médií na predaj - Regresná analýza

1. Úvod

V dnešnom digitálnom veku zohrávajú sociálne médiá kľúčovú úlohu pri ovplyvňovaní správania spotrebiteľov a predaja. Cieľom tejto analýzy je preskúmať vzťah medzi výdavkami na sociálne médiá a predajom pomocou regresnej analýzy.

2. Metodológia

Na analýzu vzťahu medzi výdavkami na sociálne médiá a predajom bol použitý dataset obsahujúci 572 záznamov. Dostupný na platforme Coursera (Google Advanced data analytics) Aplikovali sme OLS lineárny regresný model a robustný regresný model s použitím knižníc Pandas, NumPy, Matplotlib, Seaborn a Statsmodels v programovacom jazyku Python. Vzhľadom na nízku vysvetľujúcu schopnosť OLS lineárneho modelu ($R^2 = 0,281$) a prítomnosti extrémnych hodnôt v dátach sme použili aj robustnú regresiu na overenie stability a zlepšenie spoľahlivosti výsledkov.

3. Výsledky

Lineárna regresná analýza ukázala štatisticky významný pozitívny vzťah medzi výdavkami na sociálne médiá a predajom (koeficient = 14,926, $p < 0,001$), avšak model vysvetľuje len 28,1 % variability predaja, čo naznačuje, že existujú aj iné významné faktory, ktoré majú vplyv predaj. Robustná regresia potvrdila významný vplyv sociálnych médií na predaj (koeficient = 22,9648, $p < 0,001$) a poskytla vyššiu mieru istoty v odhadoch koeficientov.

Predpoklady lineárnej regresie boli testované pomocou reziduálnej analýzy. Neboli nájdené žiadne závažné porušenia. Analýza extrémnych hodnôt identifikovala štyri extrémne hodnoty v premennej „Social_Media“, ktoré by mohli skresliť výsledky štandardnej lineárnej regresie a znížiť tak presnosť modelu. Preto bola použitá robustná regresia, ktorá je menej citlivá na extrémne hodnoty a poskytuje spoľahlivejšie odhady koeficientov aj pri prítomnosti extrémnych hodnôt.

4. Zhrnutie výsledkov OLS regresie:

Tento model lineárnej regresie sa snaží predpovedať predaj (Sales) na základe výdavkov na sociálne médiá (Social_Media).

- **Závislá premenná:** Y = Predaj (Sales)
- **Nezávislá premenná:** X= Výdavky na sociálne médiá (Social_Media)
- **R-squared:** 0.281 (28,1 %)
 - To znamená, že model vysvetľuje 28,1 % variability predaja.
- **Koeficient pre Social_Media:** 21,8712
 - To znamená, že s každou jednotkou zvýšenia výdavkov na sociálne médiá sa predaj zvýši v priemere o 21,87 jednotiek.

- **P-hodnota ($P > |t|$) pre Social_Media:** 0.000
 - To znamená, že výdavky na sociálne médiá majú štatisticky významný vplyv na predaj.
- **Konštanta (const):** 122,5605
 - To znamená, že keď sú výdavky na sociálne médiá 0, predaj je v priemere 122,56.

Výdavky na sociálne médiá majú pozitívny a štatisticky významný vplyv na predaj. Model však vysvetľuje len 28,1 % variability predaja, čo naznačuje, že existujú aj ďalšie faktory, ktoré ovplyvňujú predaj.

Kľúčové body:

- Existuje vzťah medzi výdavkami na sociálne médiá a predajom.
- Vyššie výdavky na sociálne médiá vedú k vyššiemu predaju.
- Model nie je dokonalý a existujú aj iné faktory, ktoré ovplyvňujú predaj.

5. Zhrnutie výsledkov robustnej lineárnej regresie

Tento model robustnej lineárnej regresie sa snaží predpovedať predaj (Sales) na základe výdavkov na sociálne médiá (Social_Media), pričom minimalizuje vplyv odľahlých hodnôt.

- **Závislá premenná:** Predaj (Sales)
- **Nezávislá premenná:** Výdavky na sociálne médiá (Social_Media)
- **Metóda:** Iteratively Reweighted Least Squares (IRLS) s Huberovou T normou a odhadom škály Median Absolute Deviation (MAD).
- **Koeficient pre Social_Media:** 22,9648
 - To znamená, že s každou jednotkou zvýšenia výdavkov na sociálne médiá sa predaj zvýši v priemere o 22,96 jednotiek.
- **P-hodnota ($P > |z|$) pre Social_Media:** 0.000
 - To znamená, že výdavky na sociálne médiá majú štatisticky významný vplyv na predaj.
- **Konštanta (const):** 118,2877
 - To znamená, že keď sú výdavky na sociálne médiá 0, predaj je v priemere 118,29 jednotiek.

Výdavky na sociálne médiá majú pozitívny a štatisticky významný vplyv na predaj, a to aj po zohľadnení odľahlých hodnôt. V porovnaní s OLS regresiou, kde bol koeficient pre „Social_Media“ 21,87, robustná regresia odhadla tento koeficient na 22,96. Konštanta (const) bola odhadnutá na 118,29 v robustnej regresii a 122,56 v OLS regresii. Mierne rozdiely v koeficientoch naznačujú, že

odľahlé hodnoty mali určitý, ale nie dramatický vplyv na výsledky OLS regresie. Robustná regresia potvrdzuje existenciu silného pozitívneho vzťahu medzi výdavkami na sociálne médiá a predajom.

Kľúčové body:

- Existuje silný pozitívny vzťah medzi výdavkami na sociálne médiá a predajom, potvrdený robustnou regresiou.
- Vyššie výdavky na sociálne médiá vedú k vyššiemu predaju.
- Robustná regresia minimalizuje vplyv odľahlých hodnôt, čo zvyšuje spoľahlivosť výsledkov.
- Mierne rozdiely oproti OLS regresii, potvrdzujú že odľahlé hodnoty mali vplyv, ale nie dramatický.

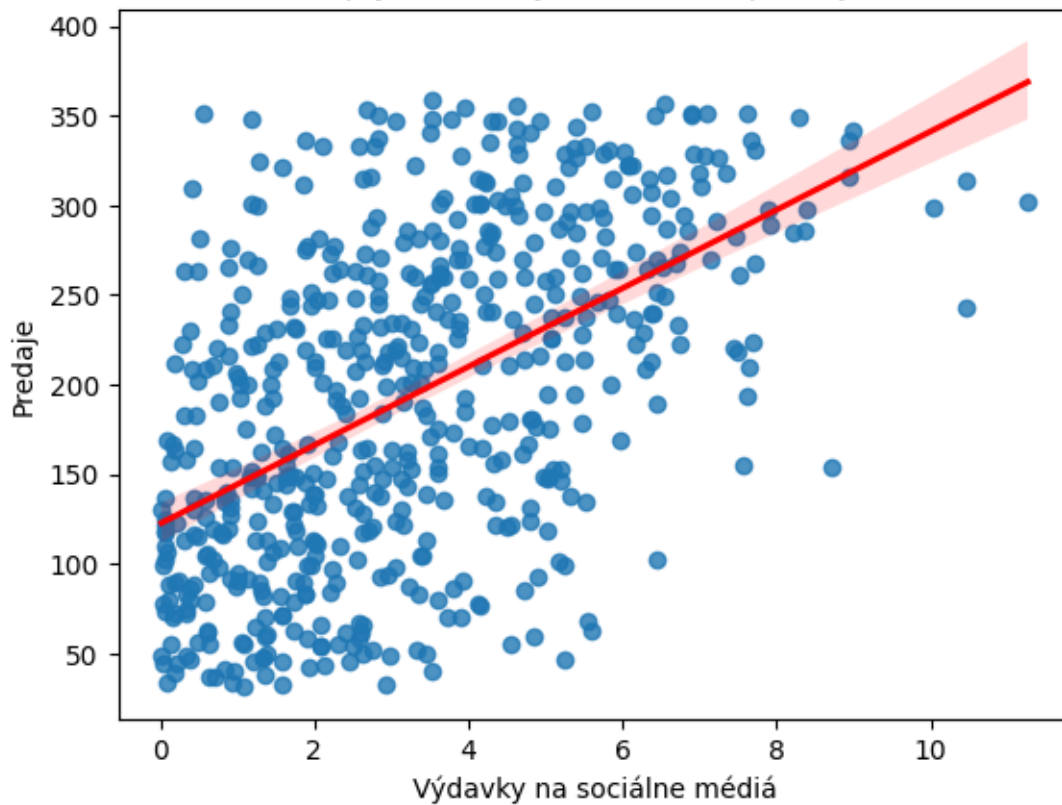
6. Záver

Tento projekt poskytuje dôkaz o významnom vplyve sociálnych médií na predaj. Zistenia naznačujú, že zvýšenie výdavkov na sociálne médiá môže viesť k zvýšeniu predaja. Slabá vysvetľujúca schopnosť lineárneho regresného modelu (28,1 %) naznačuje, že na predaj vplývajú aj iné faktory, ktoré neboli v tomto modeli zohľadnené. Robustná regresia potvrdila stabilitu a spoľahlivosť tohto záveru aj pri prítomnosti odľahlých hodnôt a poskytla presnejšie odhady vplyvu sociálnych médií na predaj.

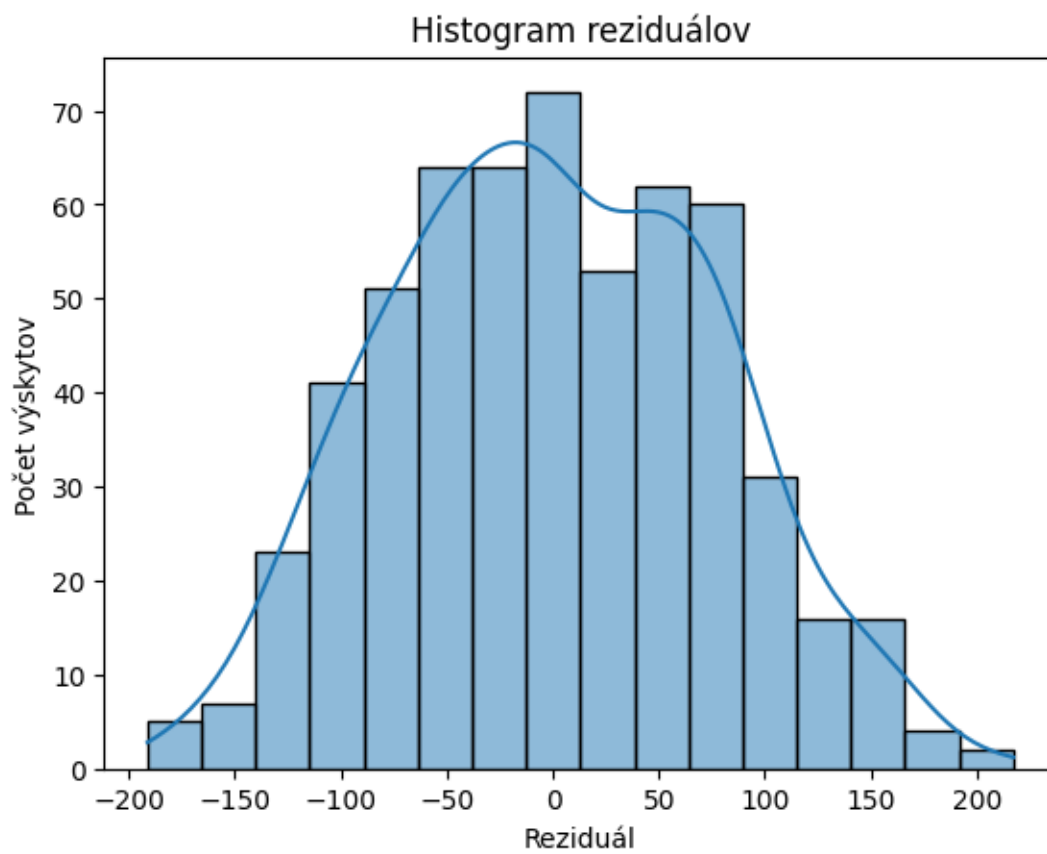
7. Vizualizácie

OLS regresia

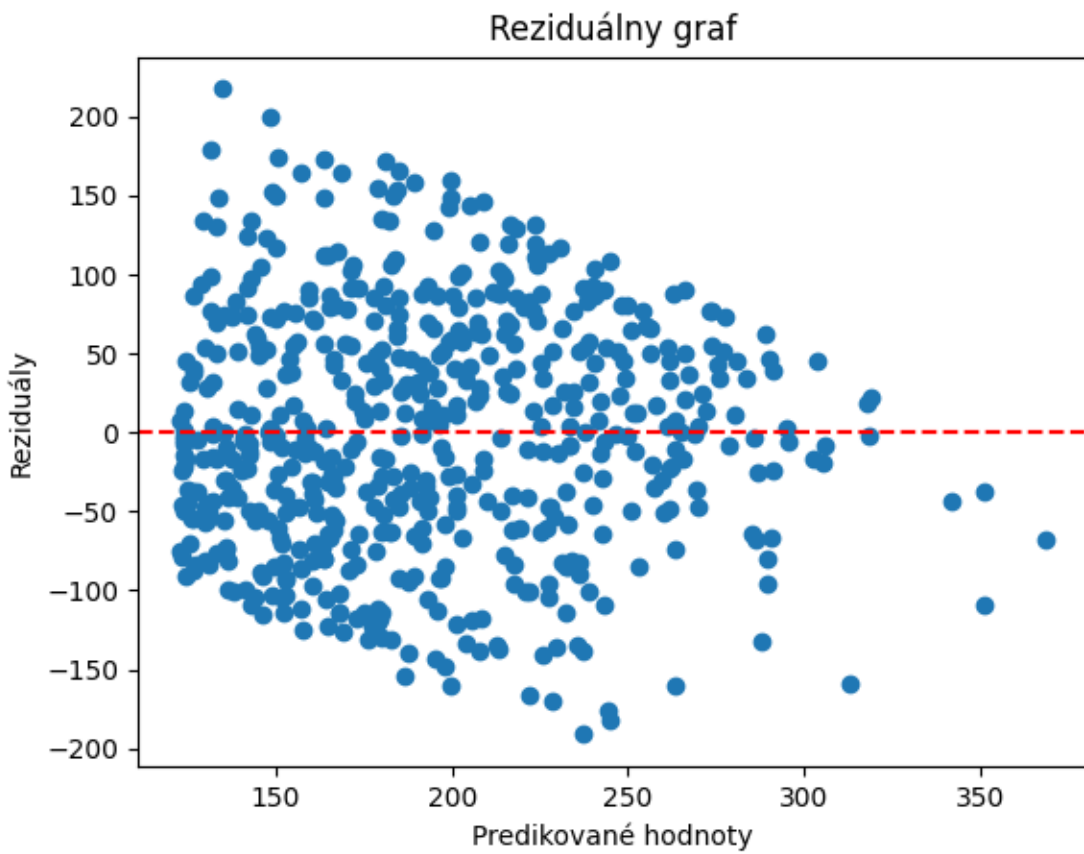
Vplyv sociálnych médií na predaje



Graf zobrazuje vzťah medzi výdavkami na sociálne médiá a predajom. Každý bod predstavuje jedno pozorovanie. Graf vizuálne ukazuje pozitívny lineárny trend, čo naznačuje, že vyššie výdavky na sociálne médiá sú spojené s vyšším predajom.



Tento graf zobrazuje histogram rozdelenia reziduálov s pridanou krivkou hustoty. Os x predstavuje hodnoty reziduálov, os y zobrazuje frekvenciu. Graf ukazuje, že reziduály sú približne normálne rozdelené so stredom okolo nuly, čo naznačuje, že model dobre predpovedá dáta.



Graf zobrazuje reziduály (rozdiely medzi skutočnými a predikovanými hodnotami) oproti predikovaným hodnotám modelu.

Robustná regresia

Graf reziduálov

