

Analýza faktorů ovlivňujících absolvování studia střední školy

Úvod

Vzdělávací výsledky žáků jsou dlouhodobě sledovány jako klíčový faktor, který může ovlivnit jejich další studijní dráhu i pravděpodobnost úspěšného dokončení školy. Předpokládá se, že lepší akademické dovednosti, jako jsou čtenářské schopnosti a matematické dovednosti, mohou přispět ke zvýšení šance na absolvování střední školy. Vedle individuálních výkonů se často zkoumá také vliv vzdělávacího prostředí, například velikosti třídy.

Cílem této analýzy bylo ověřit, zda a jak skóre z matematiky, čtení (skóre dětí ve věku devíti až čtrnácti let) a typ třídy ovlivňují pravděpodobnost úspěšného absolvování střední školy.

Popis dat

Analýza byla provedena nad daty ze studie STAR (Student/Teacher Achievement Ratio), která probíhala ve Spojených státech a sledovala vliv velikosti tříd na výsledky žáků. Data obsahují informace o 1 274 žácích, kteří byli sledováni v období přibližně od 9 do 14 let (grade 4 až grade 8).

Použité proměnné:

- **classtype** – typ třídy (regular nebo small)
- **reading** – skóre testu z anglického jazyka (čtení)
- **math** – skóre testu z matematiky
- **graduated** – zda žák později absolvoval střední školu (1 = absolvoval, 0 = neabsolvoval)

Data byla získána z materiálů ke knize Eleny Llaudetové, Data Analysis for Social Science: A Friendly and Practical Introduction na této stránce:

https://ellaudet.github.io/dss_instructor_resources/#chapter2, informace o proměnných byly získány z Harvard Dataverse, <https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/10766>.

Metodologie

Pro analýzu byla použita logistická regrese, která je vhodná pro modelování pravděpodobnosti výskytu binární události (v našem případě absolvování/neabsolvování).

Byly sestaveny dva modely:

1. **Jednoduchý model**, který predikoval absolvování pouze podle skóre z matematiky.
2. **Vícenásobný model**, který zahrnoval skóre z matematiky, skóre z čtení a typ třídy.

Výsledky logistické regrese

Jednoduchá logistická regrese

Ukázala, že skóre z matematiky má pozitivní a statisticky významný vliv na pravděpodobnost absolvování.

Každý bod navíc v matematice zvyšuje šanci na absolvování přibližně o **2,2 %**.

Vícenásobná logistická regrese

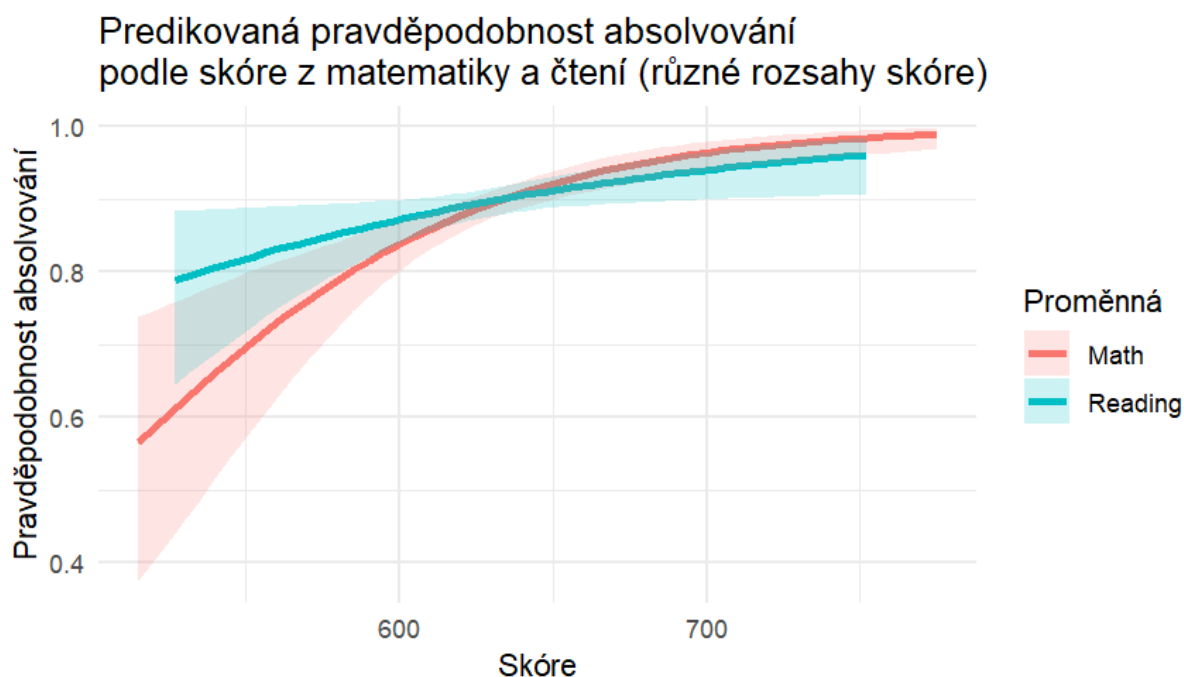
Do modelu byla přidána také skóre z čtení a typ třídy:

Matematika: každý bod navíc zvyšuje šanci na absolvování o **1,6 %** ($p < 0,001$).

Čtení: každý bod navíc zvyšuje šanci o **0,8 %** ($p = 0,0202$).

Typ třídy (small): mírně snižuje pravděpodobnost absolvování, ale tento efekt nebyl statisticky významný ($p = 0,6916$).

Vizualizace



Na grafu byly zobrazeny predikované pravděpodobnosti absolvování střední školy podle skóre z matematiky a čtení.

Obě proměnné vykazují pozitivní vliv – vyšší skóre zvyšuje pravděpodobnost úspěšného absolvování.

Vliv skóre z matematiky je patrný především u studentů s nižšími výsledky, kde rozdíly v pravděpodobnosti narůstají.

Závěr

Analýza ukázala, že vyšší výsledky z matematiky a čtení dětí od devíti do čtrnácti let mají pozitivní vliv na pravděpodobnost absolvování střední školy.

Typ třídy (small vs. regular) se jako významný faktor neprojevil.

Literatura

- Laudet, E. (n.d.). *Data Science Skills Instructor Resources*. Retrieved March 2, 2025, from https://ellaudet.github.io/dss_instructor_resources/#chapter2
- *Tennessee's Student Teacher Achievement Ratio (STAR) project*. (n.d.). Harvard Dataverse. Retrieved March 2, 2025, from <https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/10766>
- Použité balíčky v R: readr, dplyr, tidyr, ggplot2, effects