

# Faktory ovlivňující výkon studentů<sup>1</sup>

---

## Úvod

Vynikající výkon studentů je pro vzdělávací instituce často prioritní. Snahou výzkumníků je identifikovat faktory, jež mají vliv na kvalitu podaného výkonu v mezinárodně srovnávacích testech. Například podle Crosnoe, Johnsona a Eldera (2004) se mezi tyto faktory řadí vliv studenta, školy, vrstevníků a rodiny. Stejně tak podle výzkumu Hijazi a Naqvi (2006) se výkon studentů odvíjí od nejrůznějších socioekonomických, psychologických a environmentálních faktorech.



Pro účely této zprávy je využita simulovaná datová sada volně přístupné databáze<sup>2</sup>, obsahující skóre z matematiky, čtení a psaní 1000 studentů středních škol. Studenti jsou kategorizováni podle pohlaví, etnika, úrovně vzdělání rodičů, typu oběda a toho, zda student absolvoval kurz přípravy na test.

Cílem je analyzovat, které z těchto proměnných mají vliv na výkon studenta ve všech třech testových oblastech. Každá oblast je skórována na škále od 0 do 100 bodů. Je tedy třeba zkoumat tři různé modely tak, aby byly podchyceny unikátní vztahy mezi jednotlivými proměnnými a každým testem zvlášť.

## Metodologie

Pro analýzu dat je použita vícerozměrná regresní analýza. Díky modelu budeme moci měřit významnost jednotlivých proměnných, sílu jejich vlivu na výsledné skóre a porovnat je vzájemně mezi třemi modely. K odhadu regresních modelů bude použita tradiční metoda nejmenších čtverců. Jednotlivé proměnné lze otestovat pomocí T-testu. Pro určení toho, zda jsou kategoriální proměnné sdruženě významné je pak vhodné použít obecný lineární test.

Jako referenční kategorie pro etnicitu je zvolena skupina A, pro nejvyšší dosažené vzdělání je zvolena kategorie High school. Díky tomu můžeme zkoumat vliv

---

<sup>1</sup> Data a další informace o této zprávě jsou dostupné na adrese <https://dostal.vyzkum-psychologie.cz/stat4?i=278>.

<sup>2</sup> Zdroj dat dostupný zde <https://www.kaggle.com/datasets/bhavikjikadara/student-study-performance>.

ethnicity na finální skóre vzhledem ke skupině A. Zároveň také můžeme zjistit, zda vyšší dosažené vzdělání u rodičů pozitivně či negativně asociováno s vyšším skóre v testech vzhledem k referenční skupině High school.

## Příprava dat

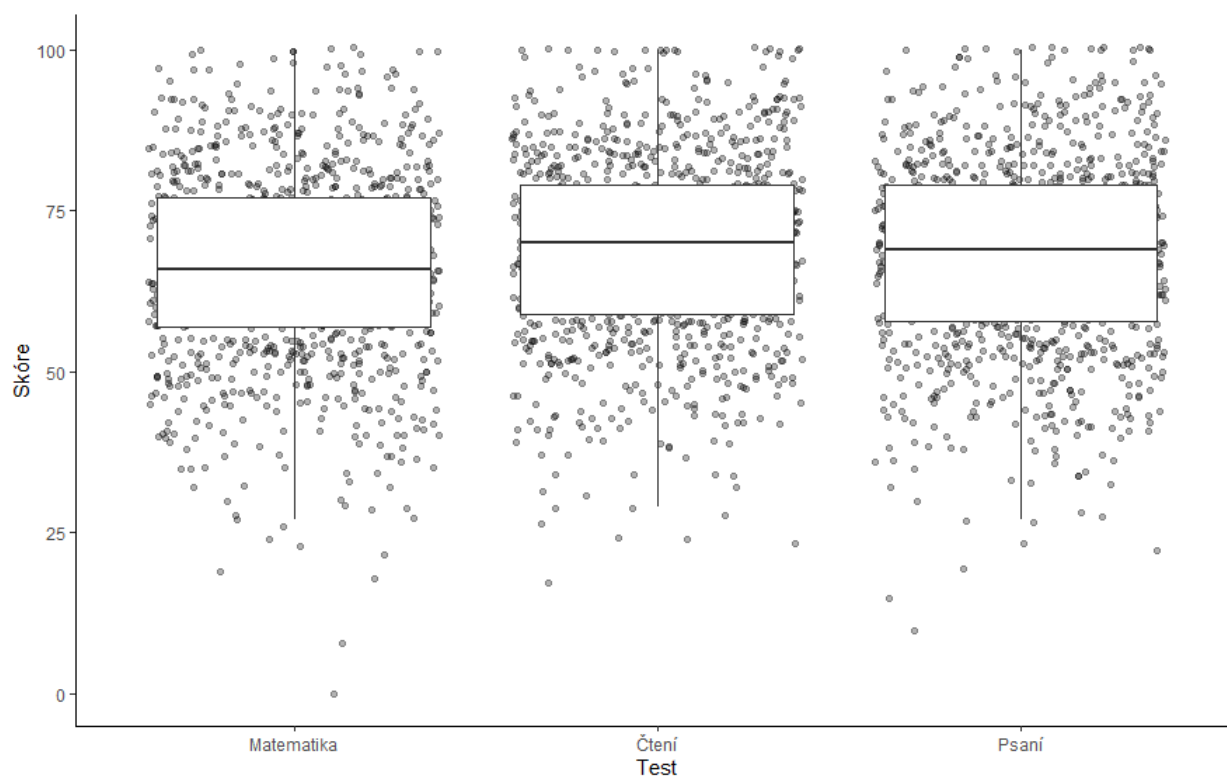
Testu se zúčastnilo 1 000 respondentů. Pohlaví, tedy proměnná *gender* byla převedena do binární proměnné *is\_male*. To samé můžeme říct o tom, zda měl student standardní oběd nebo ne a zda chodil na přípravné kurzy. Tyto proměnné jsou proto binarizovány do *standard\_lunch* a *prepared*. Etnicita studenta byla při sběru dat rozdělena do pěti skupin, které jsou v datasetu anonymizovány. Místo konkrétních etnicit tedy pracujeme se skupinami A – E. Maximálně dosažené vzdělání rodičů je v dotazníku měřeno pomocí kategorií

- some high school
- high school
- some college
- bachelor's degree
- master's degree
- associates degree

Jelikož jsou kategorie *some high school* a *some college* velmi obecné, jsou spojené do kategorie *Other*. Všechny zmíněné proměnné jsou použity jako nezávislé proměnné a budou použity k vysvětlení výsledného skóre.

Závislé proměnné, tedy skóre z testů, je v datech rozděleno do tří sloupečků. Na obrázku 1 jsou jednotlivé proměnné popsány pomocí krabičkového a bodového grafu.

Obrázek 1: Krabičkové grafy výsledků testů



Můžeme vidět, že proměnné mají velmi podobný medián, rozptyl i mezikvartilové rozpětí. Sami o sobě nám proměnné o studentech nic neřeknou a je tedy důležité vzít v potaz nezávislé proměnné.

## Výsledky

Všechny tři modely byly vytvořeny v programu R. Porovnání jejich koeficientů a p-hodnot je možné pozorovat v tabulce 01.

Tabulka 1: Porovnání regresních modelů

KOEFIČIENT	MODEL: MATEMATIKA		MODEL: ČTENÍ		MODEL: PSANÍ	
	b	p-hodnota	b	p-hodnota	b	p-hodnota
CONSTANT	47,184	0,000	58,902	0,000	55,862	0,000
IS_MALE	4,983	<b>0,000</b>	-7,081	<b>0,000</b>	-9,112	<b>0,000</b>
RACE – GROUP B	2,173	0,203	1,426	0,394	1,379	0,395
RACE – GROUP C	2,735	0,087	2,474	0,114	2,730	0,072
RACE – GROUP D	5,586	<b>0,001</b>	4,290	<b>0,008</b>	6,224	<b>0,000</b>
RACE – GROUP E	10,517	<b>0,000</b>	5,803	<b>0,002</b>	5,597	<b>0,002</b>
PARENT – OTHER	2,608	<b>0,025</b>	2,403	<b>0,035</b>	2,959	<b>0,008</b>
PARENT – BACHELOR	6,772	<b>0,000</b>	7,052	<b>0,000</b>	9,032	<b>0,000</b>
PARENT – MASTERS	7,667	<b>0,000</b>	9,087	<b>0,000</b>	10,986	<b>0,000</b>
PARENT – ASSOCIATES	4,792	<b>0,000</b>	4,892	<b>0,000</b>	5,801	<b>0,000</b>
STANDARD_LUNCH	10,850	<b>0,000</b>	7,225	<b>0,000</b>	8,170	<b>0,000</b>
PREPARED	5,344	<b>0,000</b>	7,249	<b>0,000</b>	9,878	<b>0,000</b>

Binární prediktory *is\_male*, *standard\_lunch* a *prepared* jsou ve všech modelech statisticky významné na 5% hladině významnosti. To znamená, že průměrné skóre je pro ženy, studenty se standardním obědem a studenty co absolvovali přípravný kurz vyšší, pokud jsou ostatní charakteristiky stejné.

Etnicita je ve všech modelech významná pouze jako kategorie D a E; tyto etnicity mají tedy vyšší průměrné skóre v testech než referenční etnicita A, pokud jsou všechny ostatní charakteristiky stejné. Etnicita je ve všech případech na 5% hladině významnosti sdruženě významná.

Vzdělání rodičů je na 5% hladině významné ve všech případech oproti referenční kategorii *High school*. Všechna nejvýše dosažená vzdělání mají na výsledné skóre v testu pozitivní vliv, tedy vzdělání vyšší, než high school je spojené s vyšším skóre v testu.

## Literatura

- Crosnoe, R., Johnson, M. K., & Elder, G. H. (2004). School size and the interpersonal side of education: An examination of race/ethnicity and organizational context. *Social Science Quarterly*, 85(5), 1259-1274.
- Hijazi, S.T. & Naqvi, S. (2006) Factors Affecting Students' Performance: A Case of Private Colleges. *Bangladesh e-Journal of Sociology*, 3, 1-10.