



# Analýza výsledků maturit 2018 na Gymnáziu Teplice

Jakub Mráček

## Situace k analýze, data

U maturit 2018 na Gymnáziu Teplice propadlo větší procento studentů, než bylo obvyklé (cca 14 % oproti dlouhodobě stabilním 10 %). Smyslem následujícího zadání je proto dohledat pravděpodobné příčiny náhlého zvýšení.

Vycházíme přitom z dat o známkách všech oktávánů v druhém pololetí 8. ročníku (export z Bakalářů) a o výsledcích maturit (xml export z informačního systému CERMAT). [Komplexní konsolidovaná data jsou k dispozici.](#)

## Zadání

V analýze ověřujeme hypotézu, zda vysoký počet propadlých studentů u maturity 2018 je

- přirozeným jevem,
- důsledkem náročné nebo přísně hodnocené zkoušky,
- důsledkem špatné přípravy a výbavy ze strany studentů.

Přirozeným jevem (tj. mírnou fluktuací počtu propadlých) nevysvětlíme navýšení o 50 %. Je sice teoreticky možné (v dny maturitních písemek mohl být vysoký či nízký tlak apod.), nicméně vysoce málo pravděpodobné. Druhou možnou hypotézu bychom verifikovali tak, že u statisticky významné části studentů by došlo ke zhoršení oproti dosavadním výsledkům. Poslední hypotézu potvrdíme statisticky významnou vazbou mezi výsledky zkoušky a dosavadními výsledky, ovšem tak, že nedošlo k výraznému zhoršení známky, ale tak, že prostě propadali slabí žáci.

V každé chvíli této analýzy ale musíme mít na paměti, že jde maximálně o korelace, nikoli kauzality.

## Vztah celkových studijních výsledků a výsledku maturity

Nejprve hledáme faktor, který co nejlépe koreluje s nejhorší známkou u maturity (hledáme příčiny vysokého procenta propadlíků, soustředíme se tedy na tento parametr; právě nejhorší známka u maturity je pravděpodobně to, co ovlivňuje vnímání celého výsledku - takovou hypotézu by ale bylo třeba otestovat).

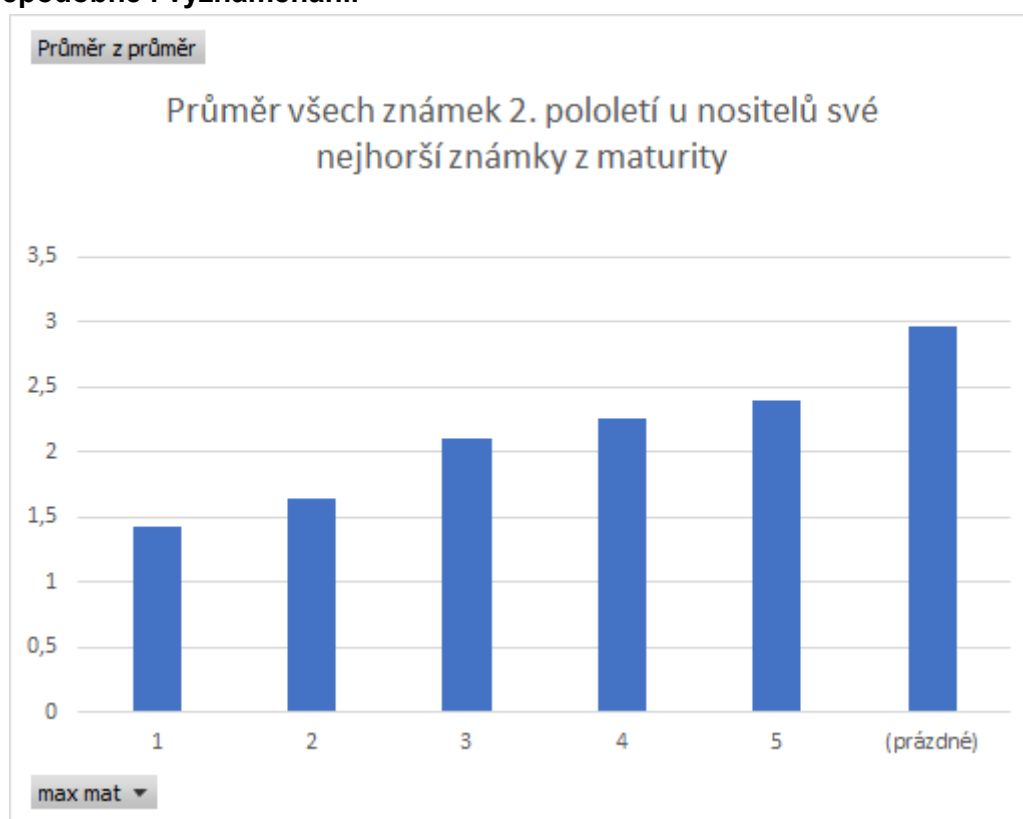
Provedli jsme tedy u všech studentů ( $N = 148$ ) výpočet Pearsonova korelačního koeficientu (PCC) s následujícími výsledky:

| Vztah   | N   | PCC   | krit. hodnota (0,01) |
|---|-----|-------|----------------------|
| Nejhorší maturitní známka/průměr všech známek 2. pololetí           | 148 | 0,496 | 0,208                |
| Nejhorší maturitní známka/hodnota 1. kvartilu známek 2. pololetí    | 148 | 0,617 | 0,208                |
| Nejhorší maturitní známka/hodnota 2. kvartilu známek 2. pololetí    | 148 | 0,429 | 0,208                |
| Nejhorší maturitní známka/hodnota 3. kvartilu známek 2. pololetí    | 148 | 0,485 | 0,208                |
| Nejhorší maturitní známka/nejhorší známka na vysvědčení 2. pololetí | 148 | 0,326 | 0,208                |

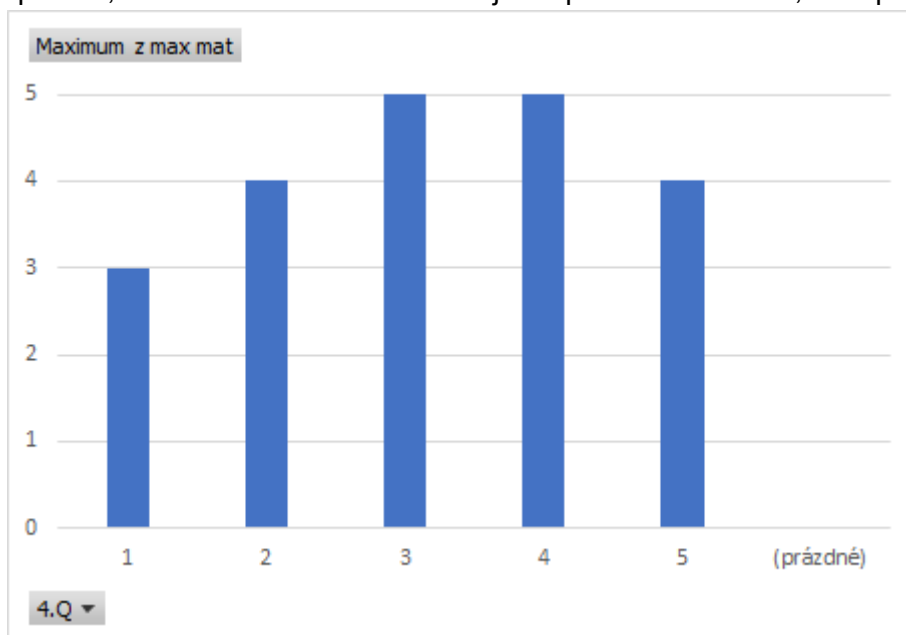
Vidíme, že korelace je na hladině významnosti 99 % značná; nejsilnější (0,496) je přitom u aritmetického průměru známek na vysvědčení. V závěsu je pak hodnota 3. kvartilu těchto známek (0,485). Jako první jsme tedy prozkoumali vztah nejhorší známky z maturitní zkoušky a průměru všech známek na vysvědčení z druhého pololetí oktávy.

Na grafu 1 vidíme průměr studentů rozdělených do 6 skupin podle nejhorší známky u maturity (skupina 1 má tedy z maturity jen jedničky, skupina 2 má alespoň jednu dvojku atd.); přičemž skupina "prázdné" ukazuje ty studenty, kteří k maturitě nešli.

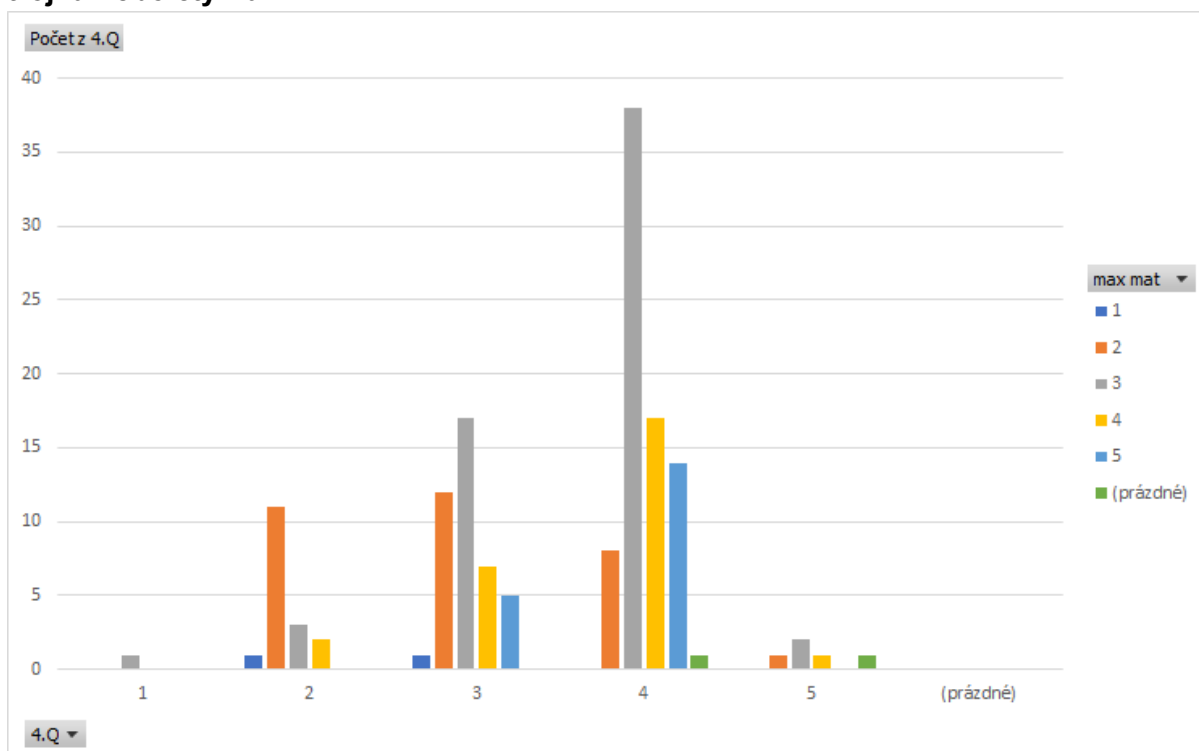
**Z rostoucího trendu je zřejmé, že k maturitě nešli studenti s průměrem blížícím se 3,0; naopak maturanti se samými jedničkami měli průměr menší než 1,5 a tedy pravděpodobně i vyznamenání.**



Graf znázorňuje nejhorší známku u maturity ve skupině studentů s nejhorší známkou na posledním vysvědčení (např. existuje student se samými jedničkami na vysvědčení, který u maturity dostal trojku). Je zřejmé, že u maturity nepropadnul nikdo, kdo měl na vysvědčení jen jedničky a dvojky, tedy vyznamenaní. Tento druhý graf tedy neguje případnou pochybnost nad grafem prvním, zda se v datech neschovávají nadprůměrní studenti, kteří propadli.



Ještě více to zpřesňuje následující graf. Studenty opět dělí podle nejhorší známky na vysvědčení a ukazuje, jakou nejhorší známku mají z maturitních zkoušek. Výše sloupce pak značí počet takových studentů. Opět tu vidíme, že student, jehož nejhorší známkou na vysvědčení byla 1 (měl tedy samé jedničky), získal u maturity trojku - jde však o jednorázový exces. **Znovu také vidíme, že propadli jen ti, kteří na vysvědčení měli alespoň jednu trojku nebo čtyřku.**



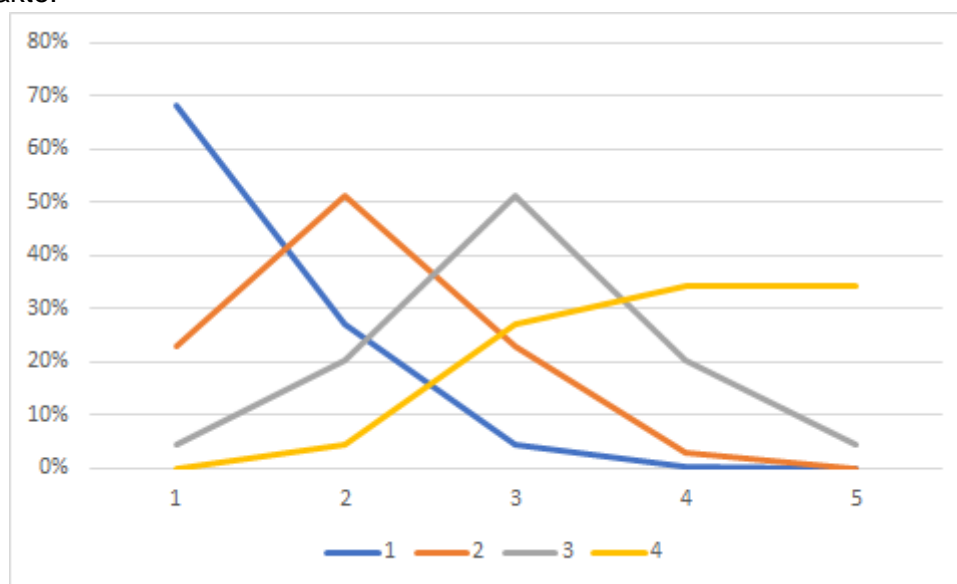
# Vztah známky na vysvědčení a známky z maturity u předmětů společné části

Podívejme se na 3 předměty společné části zkoušky ve vazbě na známky posledního vysvědčení. Na úvod opět zjistíme hodnotu Pearsonova korelačního koeficientu; i v této tabulce vidíme, že jde o silnou statistickou významnost, protože kritické hodnoty jsou výrazně (i více než dvojnásobně) překročeny.

| Vztah           | N   | PCC   | krit. hodnota (0,01) |
|-----------------|-----|-------|----------------------|
| ČJ 2P/mat spol. | 148 | 0,517 | 0,208                |
| M 2P/mat spol.  | 37  | 0,586 | 0,393                |
| AJ 2P/mat spol. | 93  | 0,510 | 0,254                |
| AJ 2P/mat prof. | 27  | 0,528 | 0,471                |

## Hypotéza

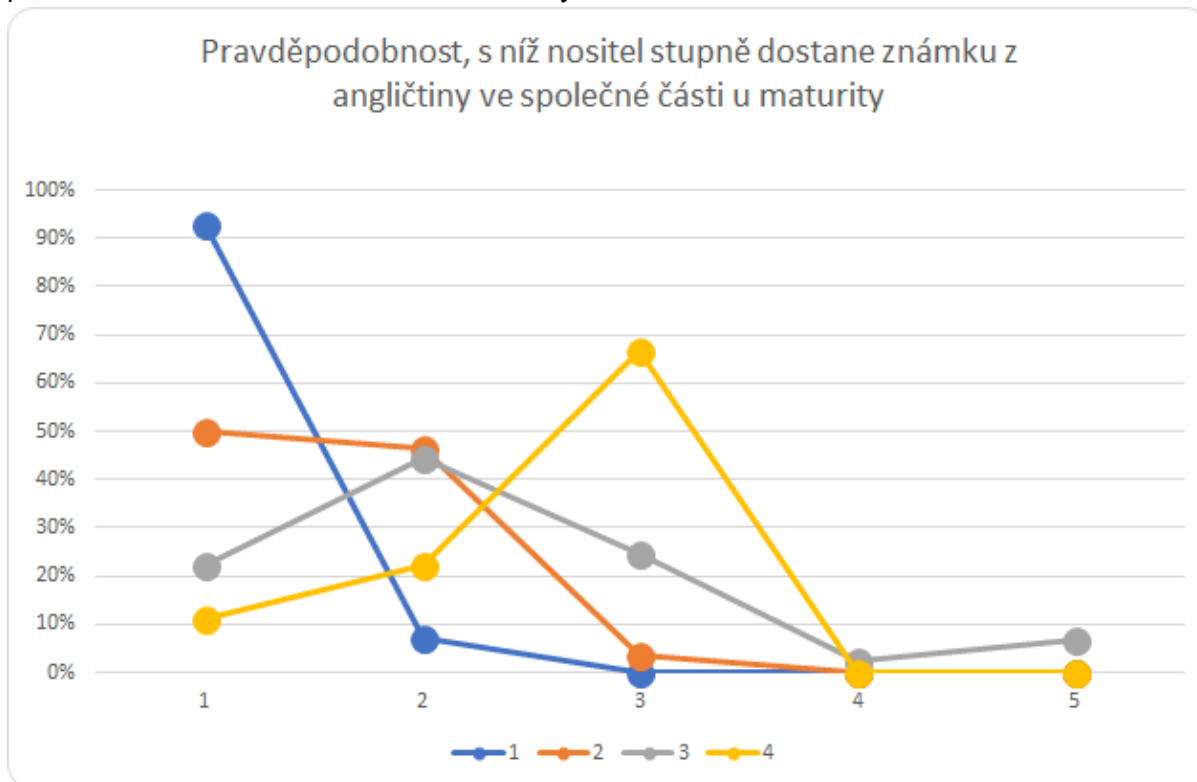
Očekáváme, že známky z maturitní zkoušky vysoce korelují se známkami ze stejných předmětů na vysvědčení proto, že se zkrátka příliš neliší - a pokud ano, pak s větší pravděpodobností o jeden známkový stupeň. Graf znázorňující distribuci známek u maturity ve vazbě na známku na vysvědčení tak principiálně vychází z normálního rozdělení a vypadá zhruba takto:



Následující grafy budeme porovnávat s tímto "etalonem" a budeme hledat odchylky od něj.

## Anglický jazyk

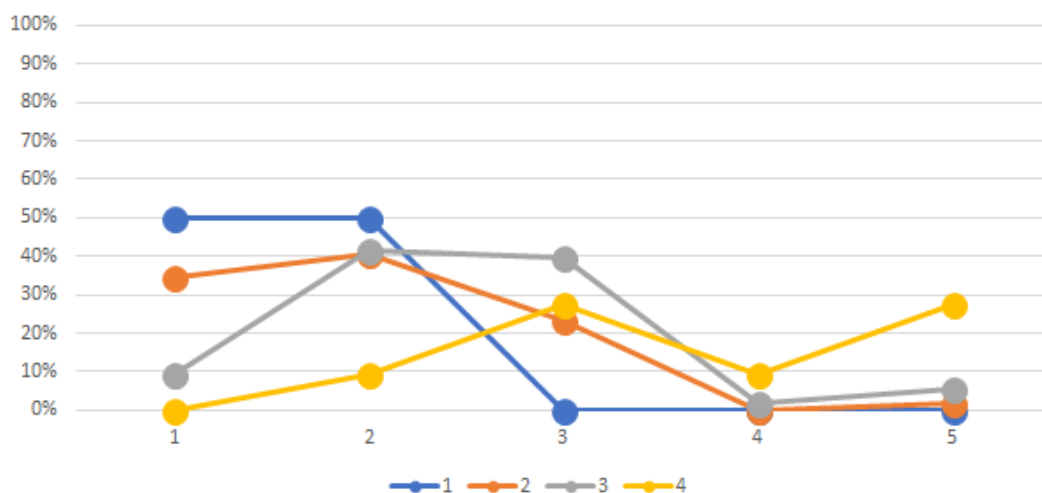
Pro státní zkoušku z angličtiny platí výrazný posun k lepšímu o jeden známkový stupeň. 93 % studentů s jedničkou na vysvědčení ji získalo i u maturity; jedničku ale také mělo 50 % dvojkařů. **Propadli 3 lidé, kteří na vysvědčení měli trojku.** Graf je až na posun doleva příkladem vzorového modelu uvedeného výše.



## Český jazyk

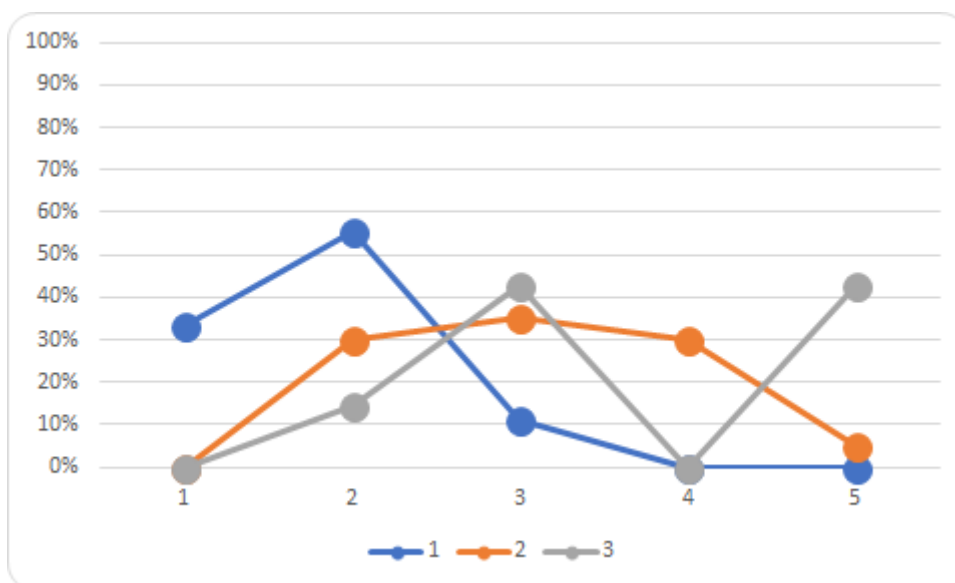
Ve státní zkoušce z českého jazyka a literatury je situace velice podobná té, kterou jsme již okomentovali u angličtiny. Překvapí nás pouze dva aspekty. Tím prvním je 100% distribuce jedničkařů mezi nositele jedničky a dvojky u maturity. **Druhou je více než čtvrtina (27 %) čtyřkařů, kteří u maturity propadli. Těm tedy maturitu z angličtiny nelze doporučit.**

### Pravděpodobnost, s níž nositel stupně dostane známku z ČJ ve společné části u maturity



## Matematika

V případě matematiky předně podotkněme, že máme k dispozici menší vzorek. I tak graf odpovídá hypotéze; oproti Čj a Aj je ale zřejmý posun vrcholu o známkový stupeň výš. Nejvíce jedničkařů tedy získalo u maturity dvojku, nejvíce dvojkařů trojku. Problémem je schod u trojkařů, z nichž vyrovnané skupiny (43 %) získalo trojku i propadlo. **Státní zkoušku z matematiky je tedy třeba trojkařům (a potenciálně též čtyřkařům) vylouvat - dokud to ještě jde.**



# Probabilistický model

Na základě výše uvedených zjištění jsme vytvořili probalisticý model známky z maturity z českého jazyka a literatury na základě nejhorší známky na posledním vysvědčení (lhostejno z jakého to je předmětu) a známky ČJ na tomto vysvědčení. Výsledky jsou přesvědčivé:

- V případech známky s pravděpodobností vyšší než 30 % je správnost predikce v 86 %;
- Bayesovská predikce pro češtinu vyšla pro nejpravděpodobnější známku ve 47 % případů, pro druhou nejpravděpodobnější ve 39 % případech případů. Probabilistický přístup tedy pro dvě nejpravděpodobnější známky platí v 86 % případů.

Z modelu a jeho otestování na stejných datech je zřejmé, že výsledky maturitní zkoušky jsou minimálně u českého jazyka dobře predikovatelné. **Pro věrohodnější ověření je však třeba modelu podrobit data z jiných let.**

## Závěr

**Analýzou jsme ukázali, že výsledky maturitních zkoušek z Čj, Aj a M ve společné části byly očekávatelné a kvantitativně (probabilisticky) jsme stanovili, kterou konkrétní skupinu s jakou pravděpodobností čeká jaká známka. Nejdůležitějším závěrem je vysoké procento propadlíků z matematiky, a to v malém procentu mezi dvojkaři a v obrovském procentu mezi trojkaři.**

**Ve vztahu k v úvodu formulovaným hypotézám je možné říct, že výsledky u maturity značně korelují s výsledky ve druhém pololetí a že šlo tedy spíše o studenty s horším prospěchem.**