

Analýza ziskovosti E-shopu

Role: Junior Datový Analytik

Cíl: Zjistit, která kategorie zboží generuje firmě největší celkový zisk.

Vítejte v realitě datové analytiky! Byli jste požádáni o vytvoření reportu ziskovosti. Data, která jste dostali z firemního systému, ale bohužel nejsou připravená na okamžité použití. Vypadají přesně tak, jak je IT oddělení vyexportovalo – surová a v různých formátech.

K dispozici máte dva zdrojové soubory:

1. **A_Prodeje.csv:** Historie objednávek (kdo, co a kolik kusů koupil).
2. **A_Cenik.csv:** Katalog produktů s nákupními a prodejními cenami.

Vaším úkolem je data načíst, vyčistit, propojit do jedné tabulky a spočítat výsledky.

ZADÁNÍ (Krok za krokem)

Krok 1: Bezpečný import dat

Načtěte oba CSV soubory do jednoho excelového sešitu (na samostatné listy s názvy *Ceník* a *Prodeje*).

Pozor: Pokud soubor jen „otevřete“ dvojklikem, data se pravděpodobně rozbijí. Soubory obsahují českou diakritiku a používají jiný oddělovač, než český Excel standardně očekává. Zvolte správný postup importu!

Krok 2: Detekce a čištění chyb (Data Wrangling)

Než začnete cokoli počítat, zkontrolujte list s Ceníkem. Najdete v něm datovou past, která vám znemožní matematické operace:

- **Nákupní cena obsahuje textové znaky**, kvůli kterým ji Excel vnímá jako slovo, nikoliv jako číslo. Tuto chybu musíte hromadně opravit.
- **Problém s desetinnými tečkami** už znáte.

Krok 3: Propojení tabulek a výpočty

Přejděte na list s Prodeji. Ke každému prodanému kusu zboží potřebujete přiřadit informace z Ceníku.

1. Pomocí vhodné vyhledávací funkce natáhněte k prodejům sloupce: **Kategorie, Nákupní cena a Prodejní cena.**
2. Vytvořte tři nové vypočítané sloupce:
 - **Tržba** (Množství * Prodejní cena)

- **Náklady** (Množství * Nákupní cena)
- **Zisk** (Tržba - Náklady)

Krok 4: Závěrečný report (Kontingenční tabulka)

Z vaší výsledné, obohacené tabulky prodeje vytvořte kontingenční tabulku.

- Zobrazte celkový **Zisk** rozpadlý podle jednotlivých **Kategorií**.
- Seřaďte kategorie sestupně (od nejziskovější po nejméně ziskovou).
- *Hotovo? Výsledek ukažte vyučujícímu ke kontrole.*

STRUČNÝ NÁVOD A NÁPOVĚDA (Když se zaseknete)

1. Jak správně importovat CSV?

Vyhnete se dvojkliku. Otevřete prázdný Excel a použijte kartu **Data → Z textu/CSV**. V náhledu zkontrolujte dvě zásadní věci:

- *Původ souboru:* Musí být **65001: Unicode (UTF-8)** (jinak se rozbije diakritika).
- *Oddělovač:* Zvolte **Čárka** (český Excel běžně čeká středník, proto by se data jinak slila do jednoho sloupce).

2. Jak na hromadné odstranění textu z čísel?

Označte problémový sloupec (**Nákupní cena**) a stiskněte spustíte funkci **Nahradit hodnotu**.

- Do pole *Najít* zadejte mezeru a text Kč, tedy " Kč" (bez uvozovek).
- Pole *Nahradit čím* nechte **úplně prázdné**.
- Klikněte na *Nahradit vše*.

Stejně můžete nahradit i desetinné tečky desetinnými čárkami.

3. Jak funguje SVYHLEDAT (VLOOKUP)?

Funkce propojující tabulky má 4 parametry:

=SVYHLEDAT(Co_hledám; Kde_to_hledám; Číslo_sloupce; Přesná_shoda)

- *Co_hledám:* Klikněte na ID Zboží v tabulce prodeje.
- *Kde_to_hledám:* Označte celou tabulku ceníku (nezapomeňte její rozsah **zafixovat klávesou F4**, aby se vám při kopírování vzorce neposouvala!), **nebo rovnou použijte její název**.
- *Číslo_sloupce:* Kolikátý sloupec z označeného ceníku chcete vrátit (Kategorie je 2. sloupec atd.).
- *Přesná shoda:* Na konec vzorce vždy napište 0 nebo NEPRAVDA.