

## Final Project

Ahoj,

omlouvám se, že tě obtěžuji, ale mám pro tebe nový úkol. Myslíme si, že objednávky, které nám poslali, jsou falešné. Nesedí nám totiž značky u některých závodníků. FIS uvádí v profilech závodníků značky lyží, bot a holí. Ne u všech, ale u některých jo. Potřebujeme ověřit, kolik objednávek je podle všeho falešných.

Prosím, zkontroluj všechny objednávky v tabulce **ski\_orders.csv** a dej mi vědět, kolik je asi falešných.

Na profily závodníků se můžeš dostat různými způsoby - například kliknutím na profil ve výsledcích nějakého podujatí, nebo otevřením URL pomocí FIS kódu. První metoda bude náročnější a méně spolehlivá, na druhou metodu mám návod:

<https://www.fis-ski.com/DB/general/athlete-biography.html?sectorcode=al&competitorid=190231>

Toto je základní URL pomocí Marca Odermatta. Tu si ale můžeš upravit a použít místo proměnné competitorid fiscode:

<https://www.fis-ski.com/DB/general/athlete-biography.html?sectorcode=al&fiscode=512269>

Takto si načteš tuto stránku. Prosím, zkontroluj to a dej mi vědět, co jsi zjistil(a).

### Úkol:

- Vytvořte si tabulku - nový dataframe, který bude obsahovat následující sloupce:
  - FIS Code
  - Jméno závodníka
  - Značku lyží
  - Značku bot
  - Značku holí
- Údaje do této tabulky získáte pomocí WebScrapingu profilů závodníků.
- Použijte již vytvořený dataframe ze souboru **ski\_orders.csv**.
- Do tohoto dataframe přidejte sloupce se značkami z profilů závodníků.
- Porovnejte značky objednaných produktů s očekávanými značkami:
  - Mějte na paměti, že v objednávkách jsou produkty jiných kategorií, takže musíte přiřadit správný sloupec, nebo kontrolu neudělat.
  - Můžete použít pomocné sloupce nebo pomocné proměnné.
  - Výsledkem by měl být sloupec, kde budou 3 hodnoty - True, False, None:
    - **True**, pokud byla objednána správná značka
    - **False**, pokud byla objednána nesprávná značka
    - **None**, pokud jde o produkty z jiných kategorií
- Vytvořte graf/grafy pomocí matplotlib:
  - Celkový podíl špatně objednaných značek
  - Podíl špatně objednaných značek pro jednotlivé kategorie

- Volba je na vás - můžete použít 4 koláčové, nebo 1 sloupcový s podílem do 100 %.

