

Praksa u kompaniji mtel

Online verzija stranice se nalazi na [ovom linku](#).

Tema: Analiza podataka

Mentor: dipl. inž. Siniša Roljić, glavni inženjer za poslovnu analitiku

Tokom trajanja prakse je planirano da se urade četiri mini-projekta:

- Projekat 1 - eksplorativna analiza podataka i upoznavanje sa SQL-om
- Projekat 2 - vizualizacija analiziranih podataka u Tableau-u
- Projekat 3 - eksplorativna analiza i čišćenje baze podataka sa SQL-om
- Projekat 4 - upoređivanje korelacije između podataka uz pomoć programskog jezika Python i biblioteka: NumPy, Pandas i Matplotlib

Bilješke | Sedmica 1

Dan 1

Datum: Aug. 16, 2021

- mentor nam je objasnio organizacionu strukturu Direkcije za tehniku kao i radne metode i uloge Službe za poslovnu analitiku u kompaniji mtel
- razgovarali smo o strukturi same prakse i zadacima koji su pred nama
- prošli smo kroz naše ranije iskustvo i tehnologije sa kojima smo radili

Dan 2

Datum: Aug. 17, 2021

- uvod u *Data Science* i *Data Analysis*
- teorijska prezentacija
- Neki od ključnih pojmova sa kojima smo se upoznali:
 - Domain Expertise
 - DWH - Data Warehousing
 - BI - Business Intelligence
 - ETL - Extract Transform Load
 - OLTP - Online transaction processing
 - OLAP - Online analytical processing
- razgovarali smo o narednim koracima i tehnologijama koje ćemo koristiti

Dan 3

Datum: Aug. 18, 2021

- počeli smo da pratimo kurs [Database Management Essentials](#)
- Neki od ključnih pojmova sa kojima smo se upoznali:
 - DBMS - Database Management System
 - Enterprise DBMSs
 - Desktop DBMSs
 - Embedded DBMSs
 - Non-procedural and procedural access
 - Transaction processing overview

- DWH processing

Dan 4

Datum: Aug. 19, 2021

- uradili smo instalaciju i podešavanje Oracle baze podataka kao i sistema za upravljanje bazama podataka Oracle SQL Developer
- prilikom instalacije Oracle baze podataka naišli smo na par problema koje smo rješavali zajedno sa mentorom
- kreirali smo jedan skup podataka kako bi se upoznali sa SQL-om
- napisali smo nekoliko SQL upita kao uvod u prvi projekat

Dan 5

Datum: Aug. 20, 2021

- prošli smo kroz [ovaj skup podataka](#)
- odredili smo koje ćemo podatke obrađivati
- uradili smo eksplorativnu analizu navedenog skupa podataka
- kreirali smo par SQL tabela za interesantnim podacima kako bi ih grafički prikazali u Tableau za drugi projekat
- planiranje narednih koraka

Bilješke | Sedmica 2

Dan 1

Datum: Aug. 23, 2021

- uradili smo vizualizaciju obrađenih podataka iz prethodne sedmice u public verziji softvera Tableau
- prošli smo kroz [ovaj skup podataka](#) za treći projekat
- uradili smo eksplorativnu analizu i formatiranje navedenog skupa podataka
- za sljedeći dan smo ostavili čišćenje podataka

Dan 2

Datum: Aug. 24, 2021

- očistili smo podatke i time završili treći projekat
 - dok smo radili ovaj projekat naišli smo na sljedeće probleme:
 - *ORA-00928: missing SELECT keyword using WITH clause*, zato što smo pokušali koristiti ključnu riječ `DELETE` nakon klauzule *WITH* što nije dozvoljeno u Oracle SQL-u tako da smo tražili alternativno rješenje
 - *ORA-01427: single-row subquery returns more than one row*, prilikom pridruživanja tabele samoj sebi, riješili smo tako što smo kreirali privremenu tabelu te smo iz nje potrebne podatke dovukli u glavnu tabelu
 - počeli smo sa radom na četvrtom projektu
 - za projekat br. 4 smo koristili [ovaj skup podataka](#)
 - koristili smo Jupyter Notebook i Python
 - tražili smo korelaciju između podataka i time smo završili četvrti projekat
-

Dan 3

Datum: Aug. 25, 2021

- kako nam je ostalo vremena krenuli smo i sa petim projektom, izvlačenje podataka sa web-a. U navedenom projektu smo izvlačili podatke o proizvodima sa Amazon-a
 - nakon što smo podatke sa sajta zapisali u kreiranu tabelu uradili smo provjeru dobijenih podataka i na osnovu cijene proizvoda smo slali email korisniku koristeći Simple Mail Transfer Protocol
 - planiranje narednih koraka i pisanje izvještaja
 - **TODO** lista:
 - ☒ dodati MARKDOWN u neke fajlove;
 - ☒ uraditi formatiranje određenog koda;
 - ☐ poboljšati grafički prikaz u Tableau;
 - ☐ pripremiti završni izvještaj
-

Dan 4

Datum: Aug. 26, 2021

- Iz prethodnog data prenosimo sljedeće zadatke:
 - ☒ pripremiti završni izvještaj
 - ☒ poboljšati grafički prikaz u Tableau;
-

Dan 5

Datum: Aug. 27, 2021

- Predavanje izvještaja
-

GitHub linkovi:

- **Source code:**
 - [Projekat 1 - eksplorativna analiza podataka](#)
 - [Projekat 2 - vizualizacija analiziranih podataka](#)
 - [Projekat 3 - eksplorativna analiza i čišćenje podataka](#)
 - [Projekat 4 - upoređivanje korelacije između podataka](#)
 - [Projekat 5 - web scraping](#)
- **Dodatak:**
 - [Skupovi podataka](#)
 - [Bilješke](#)
 - [Prezentacije](#)

Korištene tehnologije i alati:

