

Data Science képzés 1.

- Ismerd itt is körül az alkatrészt
- Data Science \rightarrow Statisztika / Prog / Üzlet \rightarrow Főleg online kurzusok
- Hadoop \rightarrow Google open source technológia
- Törzsek
- Szempontok
- Innovatív emberek

Több data \rightarrow szilárdabb hely \rightarrow sok pozíció állt rendelkezésre

Mit kell tudni?

- | | |
|--------------------------------|--|
| - Python | - Valószínűség számítás |
| - SQL \rightarrow elvileg ez | - Analízis |
| - R | - Függvények |
| - Bash | - ML modellek \rightarrow Machine Learning |

- Üzlet
- Gyakorlás
- Projekt menedzsment
- Eredményes projekt \rightarrow célszerű motiváció

1. kör \rightarrow Python és SQL

2. kör \rightarrow Statisztika (implementálás) \rightarrow Endre

3-4 \rightarrow Gyakorlás, gyakorlás

5-6 \rightarrow Szakmai pozícióra \rightarrow piaci igény \rightarrow általában lehet szoftvertanulni

Saját gyakorlás / helyi projektek! \rightarrow fontos! \rightarrow tanulni lehet belülről és jó is portfóliónak

API \rightarrow ez lehet szilárd az adatokhoz

- Alkalmazás -> de az adatokat azaz

Egy projekt lépése:

1. Adatgyűjtés -> minden kell adatokból -> a felhasználókat kettőtől kezdve gyűjtik

2. Adatátvitel -> pl. txt fájl

első dimenzió és egy többdimenziós táblázat (SQL)

3. Adatintegráció

4. Adatellenőrzés -> data egészen csinosítja fel
jó dolog, jó válaszok

Kéne mindig emlékezni arra, hogy
mi az a cél.

5. Visualizáció -> egyszerűen lehet kommunikálni

6. Döntés -> adat -> allokáció

SQL alapok:

- SQL leírás

- azaz a kódok

- gyorsabb az SQL, mint az Excel

- Könnyebb tárolni az adatokat

- PostgreSQL vs Google Analytics

↓
de az adatokat definiálni a rendszer

- MySQL

- MS SQL -> kicsi költséggel kezel

- SQLite

- Valamikor nem jó az SQL -> pl. predikciók

- Képesek tanulásra

- gyakorlat -> Halli projekt!

- Stackoverflow → gyors és minden fort van

SQL a gyakorlaton

- data.com/sql-day-101/

- SELECT * FROM table_name; → lekérdezés

↓ ↓ ↓
típusok tábla tábla neve

- Integer → egész szám : 5, 20, 0, -20

- Decimal → nem egész számok 20.01, 3.1415, 20.000

- Boolean → igen / nem ⇒ True / false / 0 / 1 / on / off / yes / no

- Text → szöveg → 'R2D2', 'Holt'

- Date → dátum → 2017-12-31 → ezt lehet állítani

- Time → idő → 23:59:59

- Database Explorer → feltehetően az adatbázis

Data file → ott a tábla

- Select animal, unique_id, water_need, animal FROM zoo;
⇒ Oszlopokat is lehet kiírni

- Select * FROM accidents limit 10;

↓ ↓
accidents első 10 sor

- Időnként sok → így megy mindig

- Select city, accident_date, damage FROM accidents;

- Select * FROM zoo Where animal = 'tiger';

↓ ↓ ↓ ↓
választ tábla feltétel/alkalmazás szöveg

- Datamod is related "jelle"

Where water_need > 500;

↓

Comparison operator → '='

<>

!=

<=

- Select * From accidents Where damage >= 1;

→ Format SQL parameter names

- Select * From zoo Where animal Like '%o%';

↓

to know what is the letter 'o'

'%o%g%r%'

- Select * From accidents Where city Like '%Alt%';

⇒ Tell degree is bit hot or not

- Select *

From zoo

Where animal = 'zebra'

And/OR water_need > 500;

↓

is / copy

- Select *

From zoo

Where animal

water_need Between 100 And 500;

- Select city, country, accident_date, damage

From accidents

Where damage < 100

OR damage > 1000

Order By → növény sorrend

ASC → növekvő
DESC → csökkenő

Select *

From zoo

Order By animal Desc, water_need ASC;
↓
növény

Select Distinct (animal) → eltünteti a duplikációkat

From zoo;

- Select * / reported → csak a legújabb

From accidents

Order By damage Desc

limit 3;

- Select *

From accidents

Where city = 'Barcelona'

Order By damage Desc

limit 3;

- Select *

From accidents

Where (city like 'Z%' OR city like 'MAGD')

And accidents date '2007-01-01'

Order By damage DESC

limit 10;

Select

Min (water_need),

Max (water_need),

Count (water_need),

Sum (water_need)

Select Count (Distinct (city))

From accidents;

= - => Egy komment

-- Itt egy átlagot számolt => (átlagos) kommentek

SQL függvények + Group by

-> szintaxis = 'ed'

Select animal, AVG(water-used) Sum(water-used)

From zoo

Group By animal;

- elephant (1,5) [↑] his csoport

Select car_brand, Sum(damage)

From accidents [→] ez az a csoport animal

Group By car_brand;

AS -> alias -> átnevezés az az a csoport

Subquery -> egy másik lekérdezés a lekérdezésben

Select country * AVG(damage) AS avg-damage

From accidents

Where weather = 'Rainy'

OR weather = 'Windy'

Group By country;

Workshop

Send & Tree

my-date, my-time, user-id, device-type

country

device type

country source

Workshop

Send to tree

my_date, my_time, user_id, device_type

country

device_type

country_code