PA220: Database systems for data analytics

Home Assignment 4

Vlastislav Dohnal

HA 4: Create reports in Power BI

- Install [Power BI Desktop](#) (it's free)
    - [Tutorials by MS](#), [Jak na Power BI](#), [MS Showcase on YouTube](#)
- Tasks:
    1. Import your Pg tables from Stratus VM
    2. Create reports (in detail in the next slides)
        - Page 1
            - Number of connections per country & slicer on time of connection
        - Page 2
            - Changes in traffic from September to October
        - Page 3
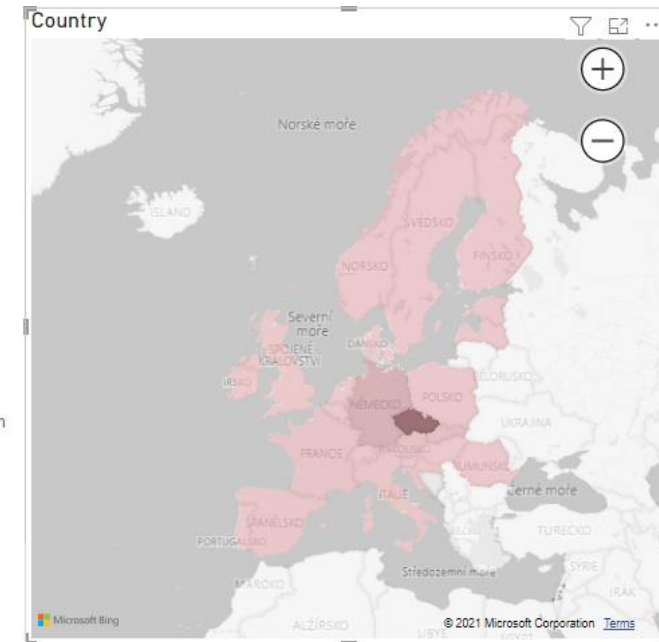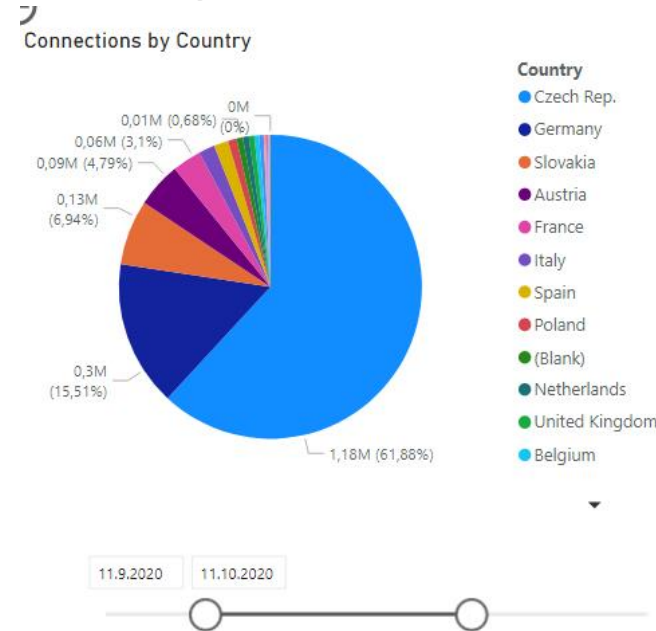   - Compare traffic during a public holiday to a regular day

HA 4: Create reports in Power BI - Tasks

- Tables (2 pts.)
  1. Import your DW's Pg tables (*import*, not direct query)
     - If not possible to load all data, take a subset of the fact table (min. September and October of year 2020)
  2. Check
     1. data type of individual columns
        - you may need to remove some imported columns and add new ones (aka their duplicates), e.g., day of week derived from timestamp and, …
     2. existence of relationships from dimension tables to the fact table
        - It should be 1:N.
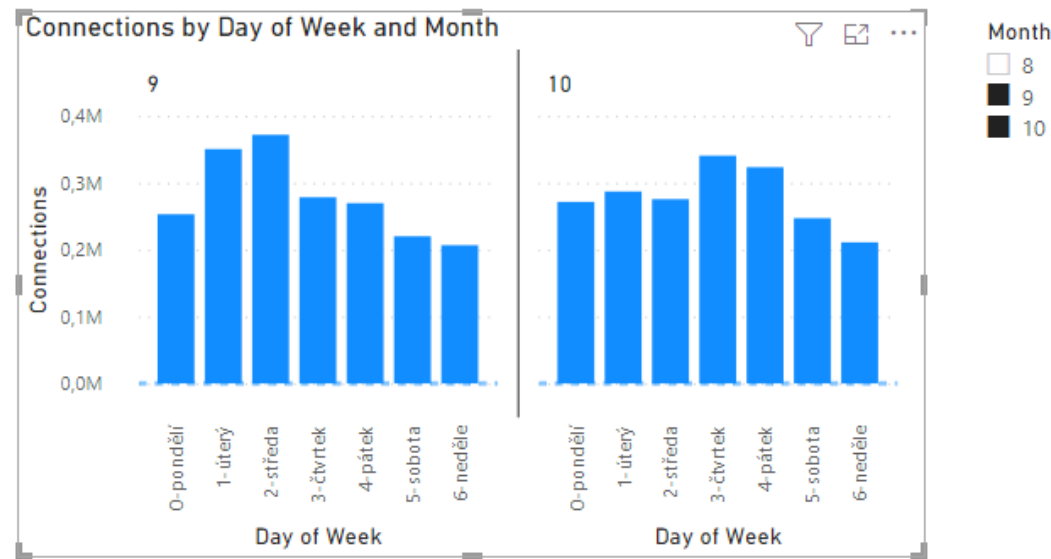
HA 4: Create reports in Power BI - Tasks

- Page 1 (4 pts.)
  - Create a **chart** for the number of connections per country of data transfer,
  - and display a **map**
    - Filling color intensity will respect the amount of connection in the country
  - Add a data slicer for connection time
    - Pre-set the slicer to limits from Sep. X to Oct. X, 2020, where X is the day of your birthday
      - e.g., X=11 if your birthday is on February 11.
    - Then make the screenshot.

- E.g.,

HA 4: Create reports in Power BI - Tasks

- Page 2 (2 pts.) – Changing traffic from month to month
  - Create a bar chart for the number of connections per day of the week, separately for individual months
    - At least, you must compare September to October in 2020.
    - E.g.,

HA 4: Create reports in Power BI - Tasks

- Page 3 (2 pts.) – Compare traffic during a public holiday to a regular day
  - Compare the traffic during Wednesday, Oct 28, 2020 with Wednesday, Oct 21, 2020:
    - in number of connections per country as well as the total number of connections.
    - display it as a table
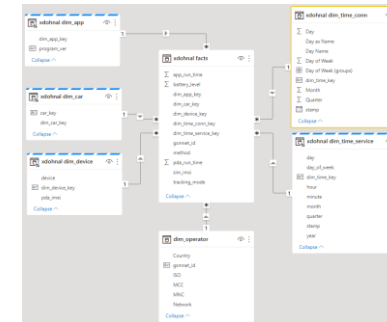   - E.g.,

Connections on Oct 28		Connections on Oct 21	
Country	Count	Country	Count
Czech Rep.	29491	Czech Rep.	36119
Germany	16809	Germany	17528
Austria	4686	Austria	4745
Slovakia	3542	Slovakia	3577
France	3363	France	2951
Italy	1475	Italy	1149
Spain	1027	Belgium	971
United Kingdom	836	Poland	947
Hungary	727	United Kingdom	910
Netherlands	539	Spain	855
Poland	372	Hungary	503
Belgium	369		406
Sweden	230	Netherlands	356
Denmark	166	Sweden	183
	165	Switzerland	112
Norway	82	Slovenia	66
Portugal	71	Serbia	62
Switzerland	65	Luxembourg	49
Luxembourg	56	Norway	38
Total	**64071**	Croatia	35
		Denmark	6
		Total	**71568**

Submission of Assignment 4

- Hand into the IS vault:
  - a single ZIP archive containing:
    - `reports.pbix` – Power BI project
    - `model.png` – screenshot of your model
    - `page1.png` – screenshot of Page 1
    - `page2.png` – screenshot of Page 2
    - `page3.png` – screenshot of Page 3

- Grading
  - total 10 pts
  - see previous slides for break-ups

Importing DW data

- Data source is "PostgreSQL database"
  - Mind the faculty's firewall:
    - use FI VPN and then PuTTY to log in into your Stratus VM
      - use the private key to authenticate,
      - forward local port 5432 to locahost:5432

- You may use direct query (leave data in DB)

- or import (copy data locally).
  - this is preferred

Choose tables

E.g.,

- Filter by your login

- Select DW tables & load
  - If time dim is referenced multiple times, you will add the next table later by analogy.