Data Analytics in Health Care and Connected care: project

TAS: Joris Wuts,, Redona Brahimetaj

March 2023
Prof. Bart Jansen, Prof. Jef Vandemeulebroucke, Prof. Nikolaos Deligiannis

VRIJE
UNIVERSITEIT
BRUSSEL

Overview

- Getting the data on Bigquery

- Bigquery

- Usefull links

Getting the data on Bigquery

1. Send request to physionet after completing the CITI test
2. Go to the mimic iii website on [physionet](#)
3. Add your gmail account to your account email addresses and add couple it to cloud computing in physionet
4. Confirm your gmail account in your mailbox
5. Go back to the mimic iii page and scroll down to the data section
6. Click on 'request access using Bigquery'
7. Your page should refresh with a green bar in the top confirming your access
8. Open [gcloud console](#)
9. Create a free project and give it a name: Project-mimic
10. Go to bigquery scrolling in the dashboard on the left
11. Click add dataset by name and enter : physionet-data
12. You now have access to the mimic-iii dataset in bigquery

VRIJE
UNIVERSITEIT
BRUSSEL

Bigquery

- What is bigquery ?
  - Part of google cloud platform
  - Managed data warehouse
  - Scalable
  - Querying via SQL
- Tips and Limits:
  - Max 1TB per user per month of processed queries
  - Only use the graphical interface of bigquerry to process data or colab
  - Download csv files with the output from GCP
- SQL tips
  - Inspect your tables and know your data
  - Never do SELECT *, but identify the necessary fields only
  - Set LIMIT to the sql output in the query during development
  - Avoid 'DISTINCT'
  - Use dry runs to estimate the processed bytes

VRIJE
UNIVERSITEIT
BRUSSEL

Useful links

- [Colab](#)
- [Mimic iii example queries](#)
- [Physionet credentialing](#)
- [Physionet mimic iii access page](#)

VRIJE
UNIVERSITEIT
BRUSSEL