Friedrich-Alexander-Universität
Technische Fakultät

FAU

Physical Activity and Health in Europe

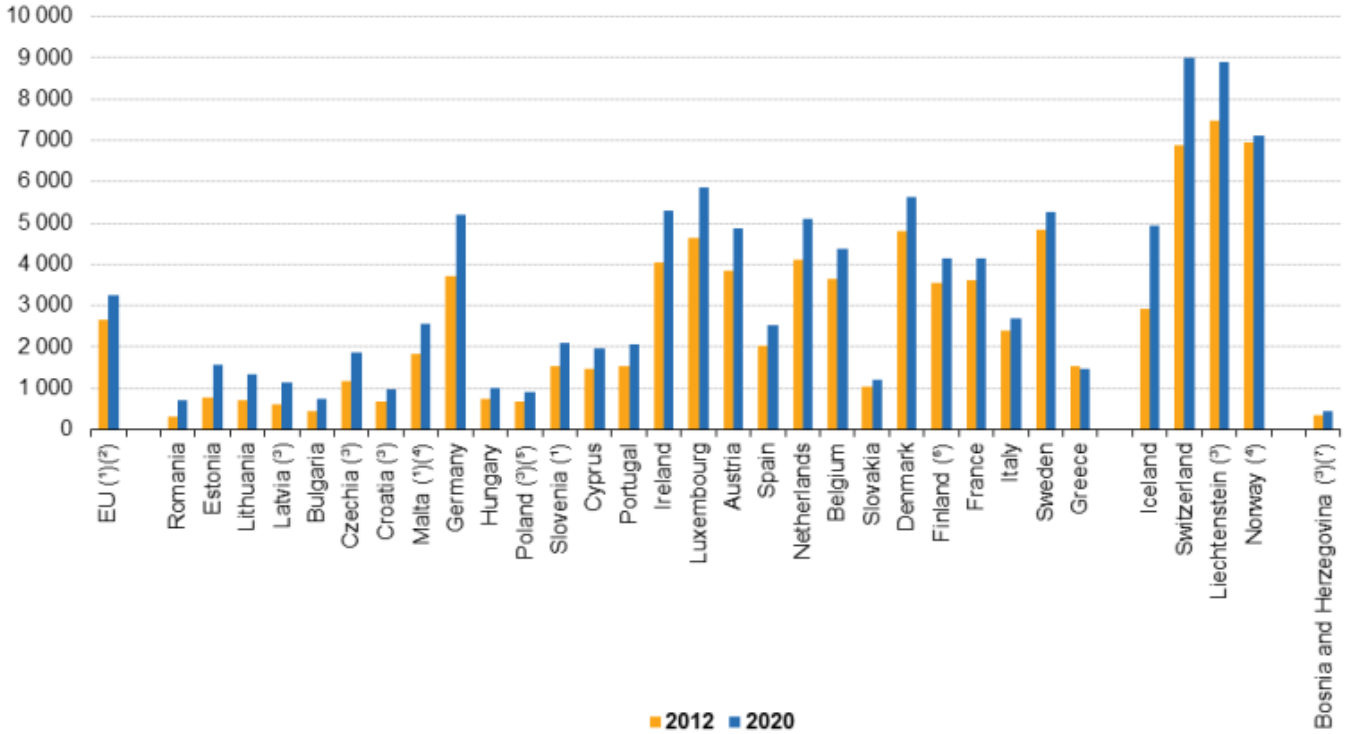
Leonie Färber
Methods of Advanced Data Engineering
WS 23/24

Motivation

Motivation

Current healthcare expenditure per inhabitant, 2012 and 2020

(€)



Note: ranked on the overall percentage change between 2012 and 2020.

(1) 2014 instead of 2012.
(2) 2020: estimate.
(3) 2013 instead of 2012.
(4) 2019 instead of 2020.
(5) 2020: provisional.
(6) Break in series.
(7) 2018 instead of 2020.

Source: Eurostat (online data code: hlth_sha11_hf)

eurostat

Motivation

- Economic benefits (e.g. reduced healthcare costs)
- Disease prevention
- Sedentary behaviours increase
- Research suggests impact of physical activity on immune system and QoL

→ Can we find a clear correlation between activity levels and health in Europe?

Methods

Methods

Data Sources

- Physical activity
  - Data type: CSV
  - Source: Eurostat
  - License: CC BY 4.0
- Depressive symptoms
  - Data type: CSV
  - Source: Eurostat
  - License: CC BY 4.0
- Self-report on general health
  - Data type: CSV
  - Source: Eurostat
  - License: CC BY 4.0

Methods

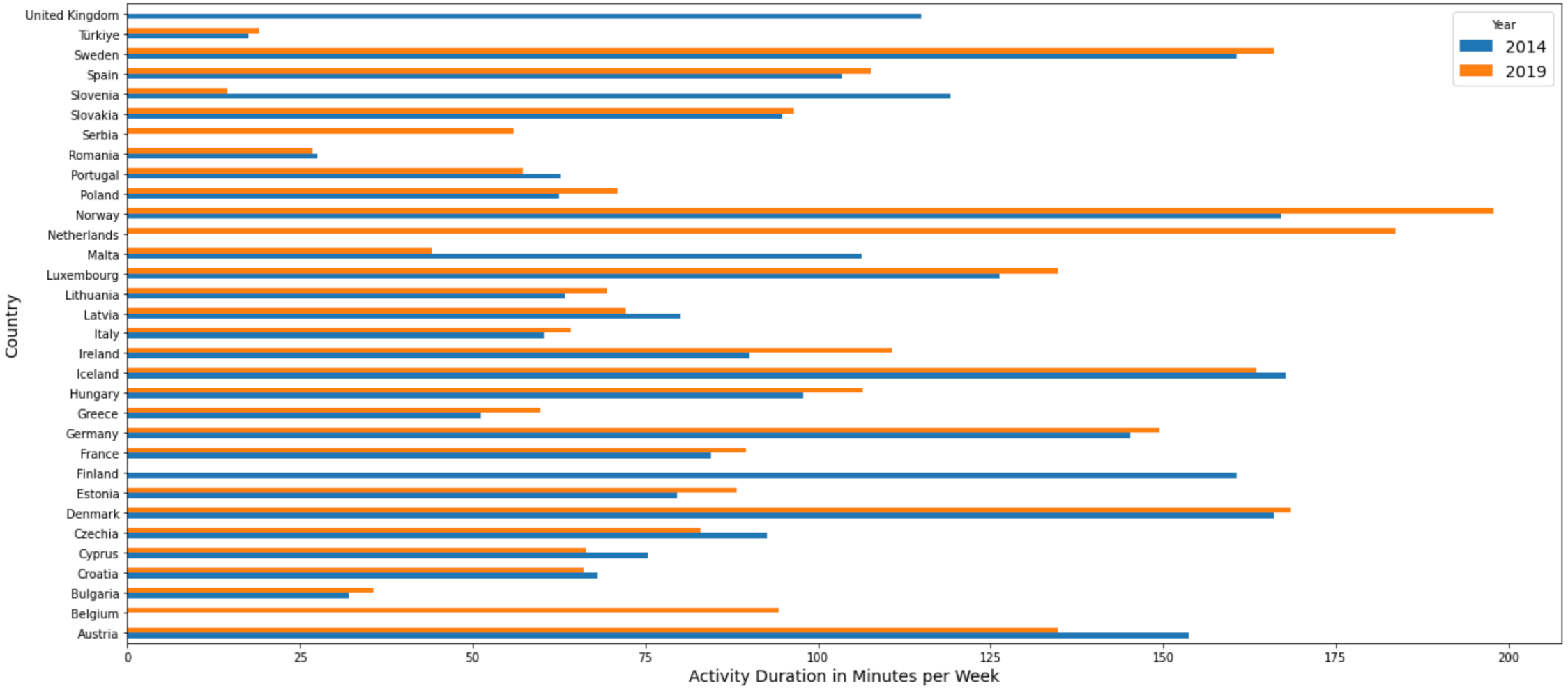
Data Pipeline

- Extract
- Validate
  - Generalized values (e.g. all of Europe)
  - Standardize values (e.g. age = “total”)
  - Overlapping intervals
  - Observation flag
- Reshape
  - Standardized columns
  - Columns containing metadata
- Transform
  - Transform string intervals to numbers
- Load

Results

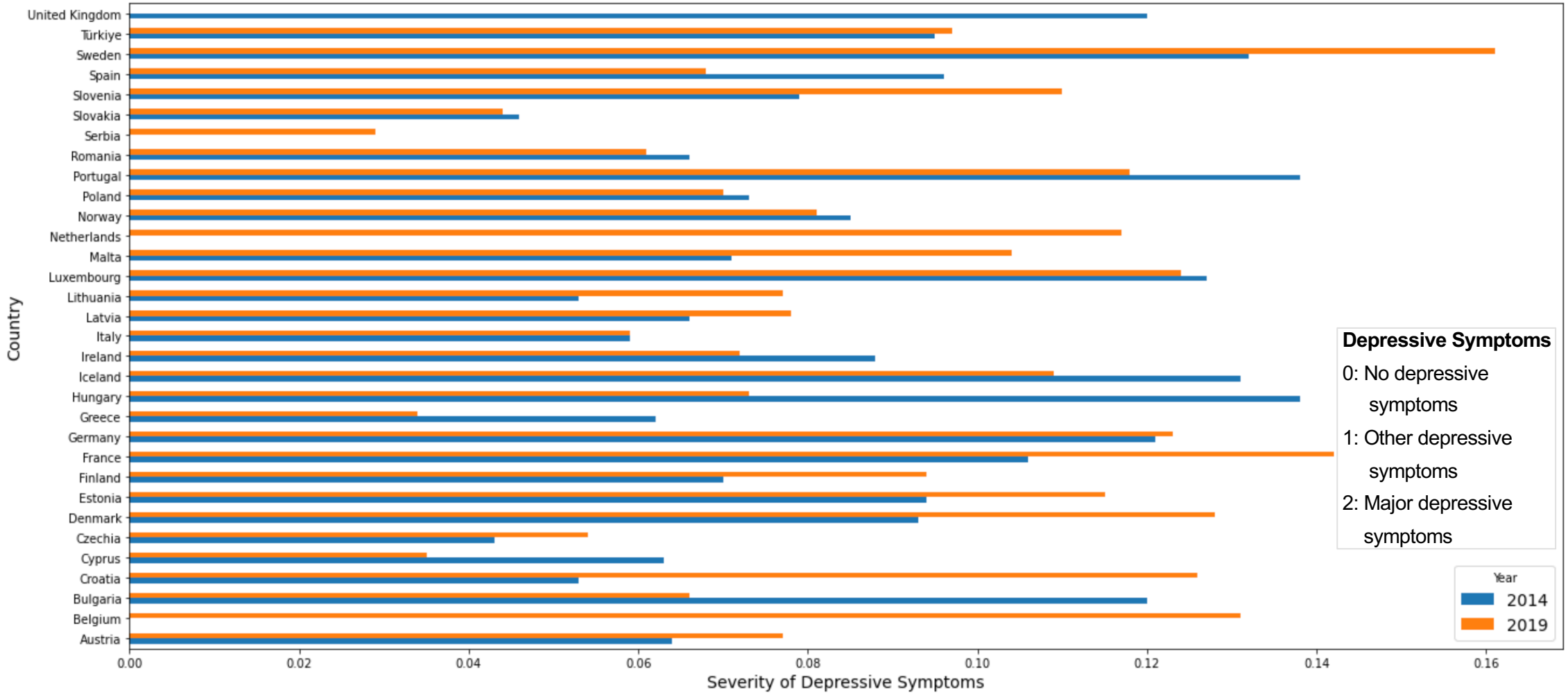
Results

Data Presentation



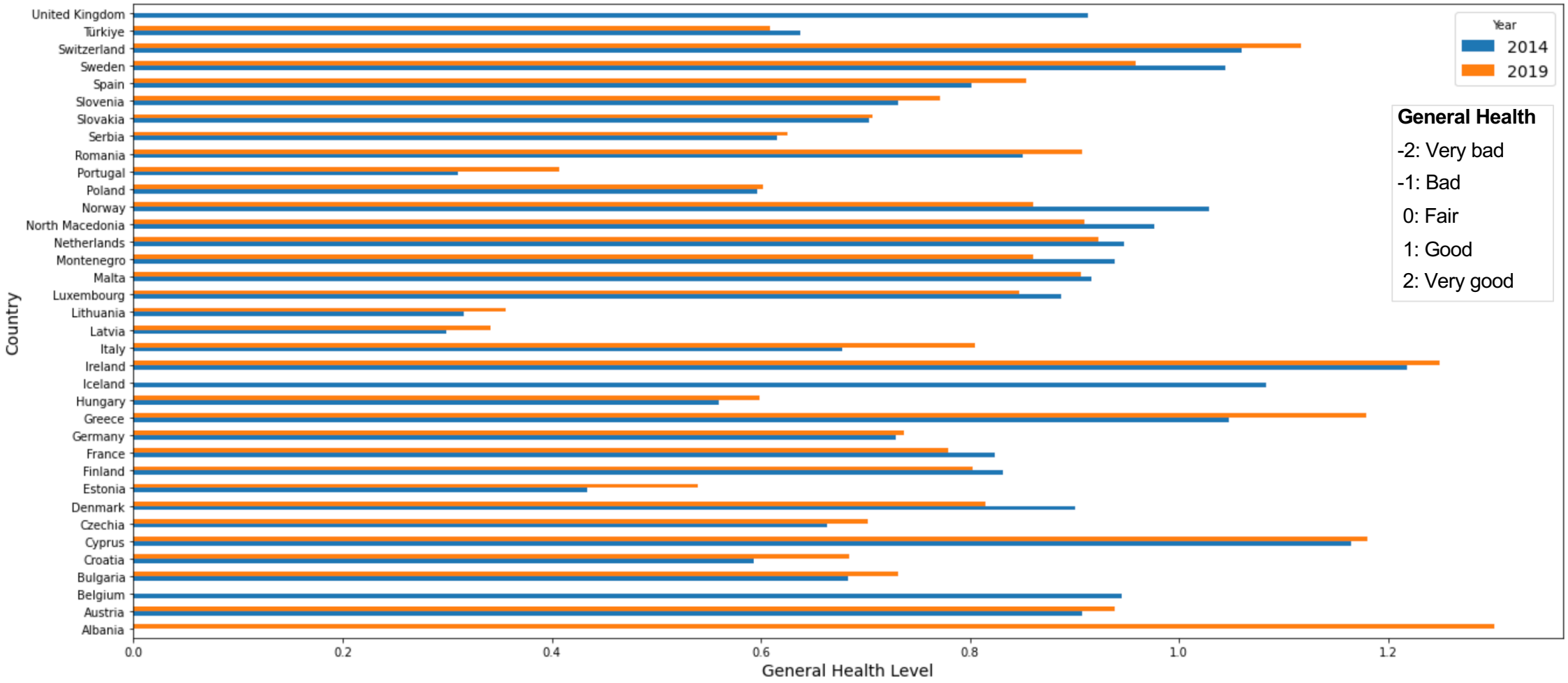
Results

Data Presentation



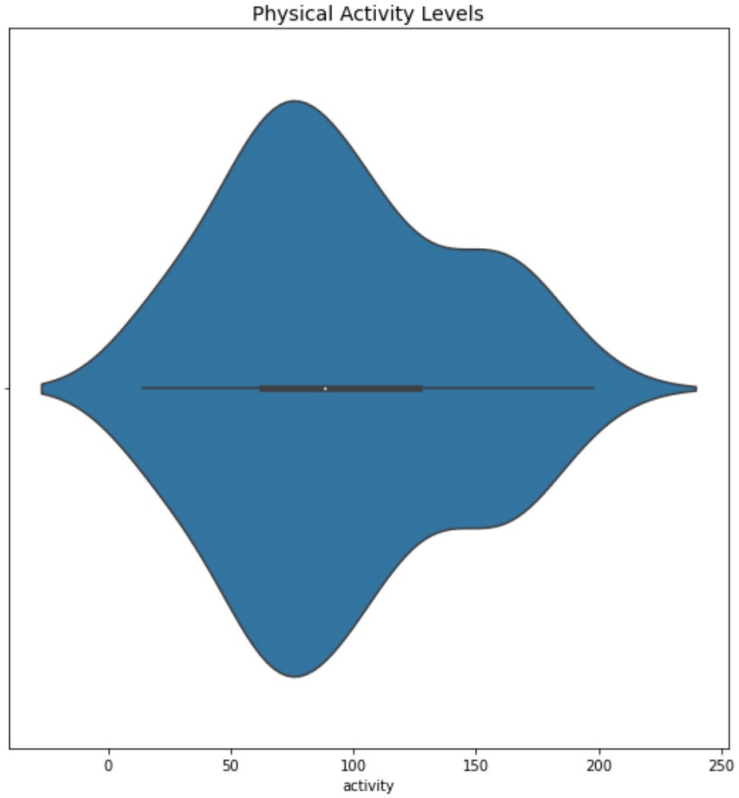
Results

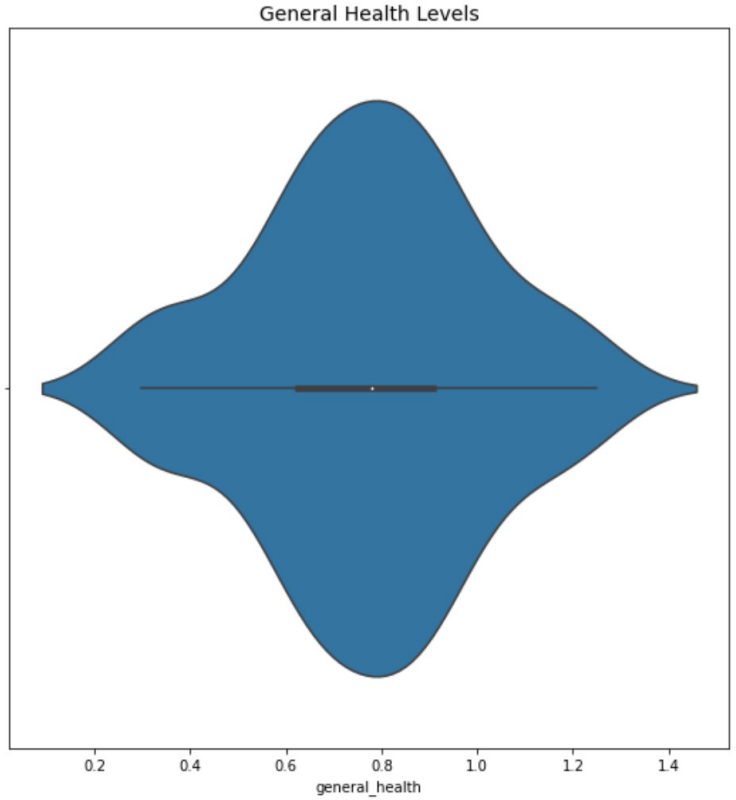
Data Presentation

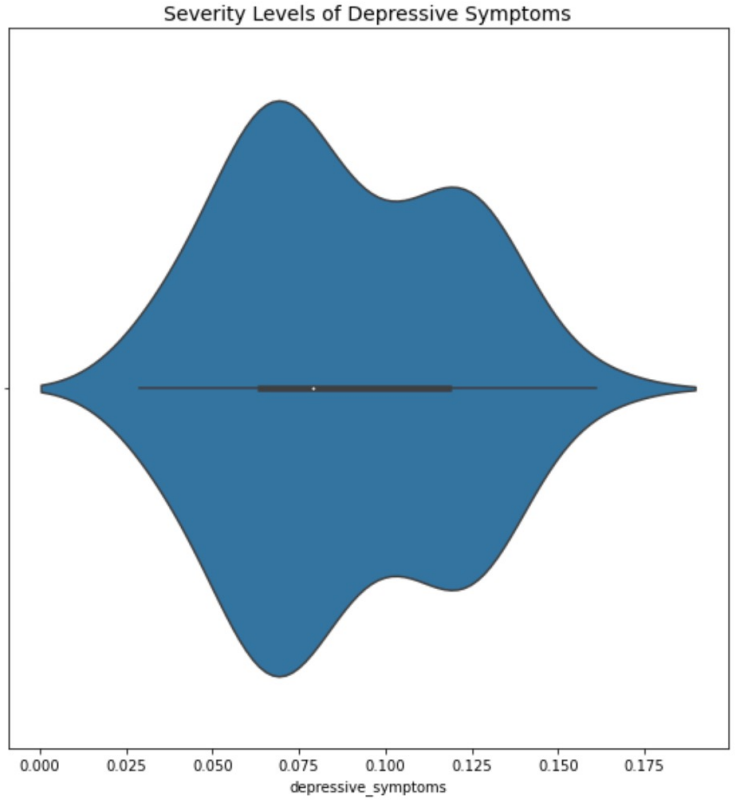


Results

Data Distribution

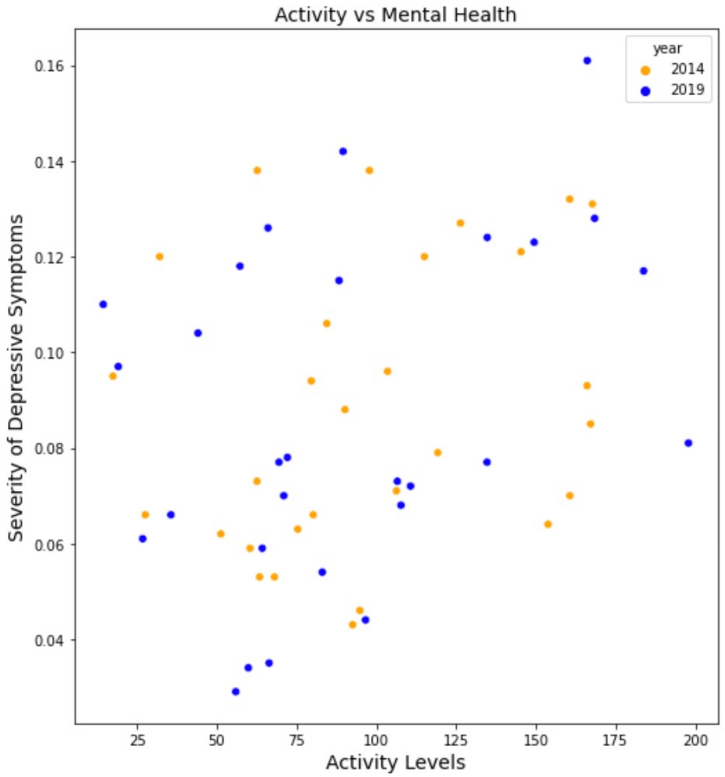


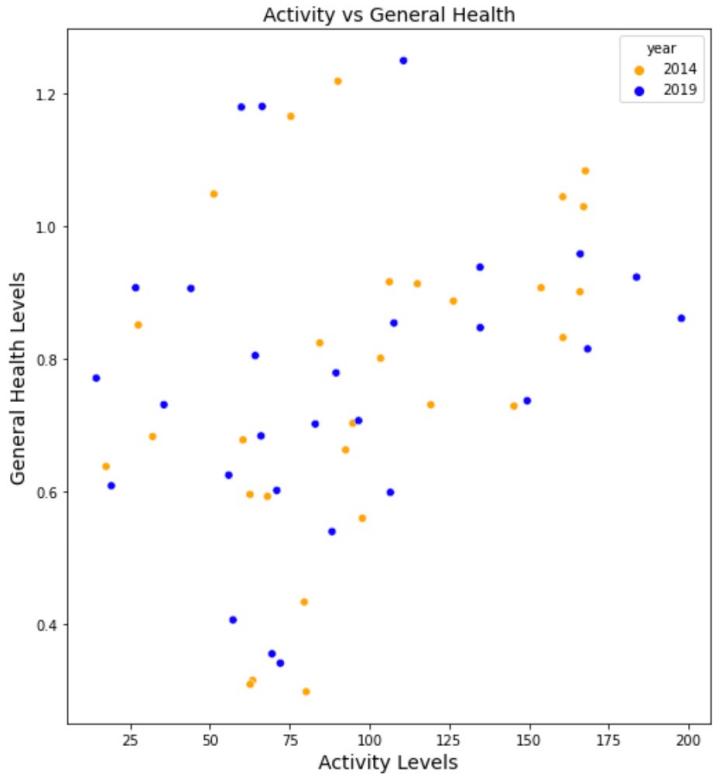


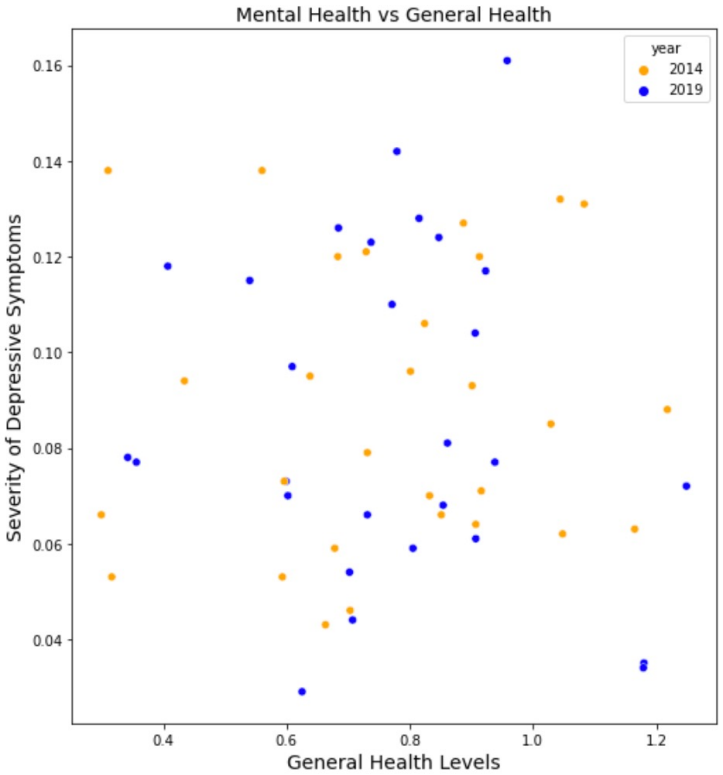


Results

Data Correlation

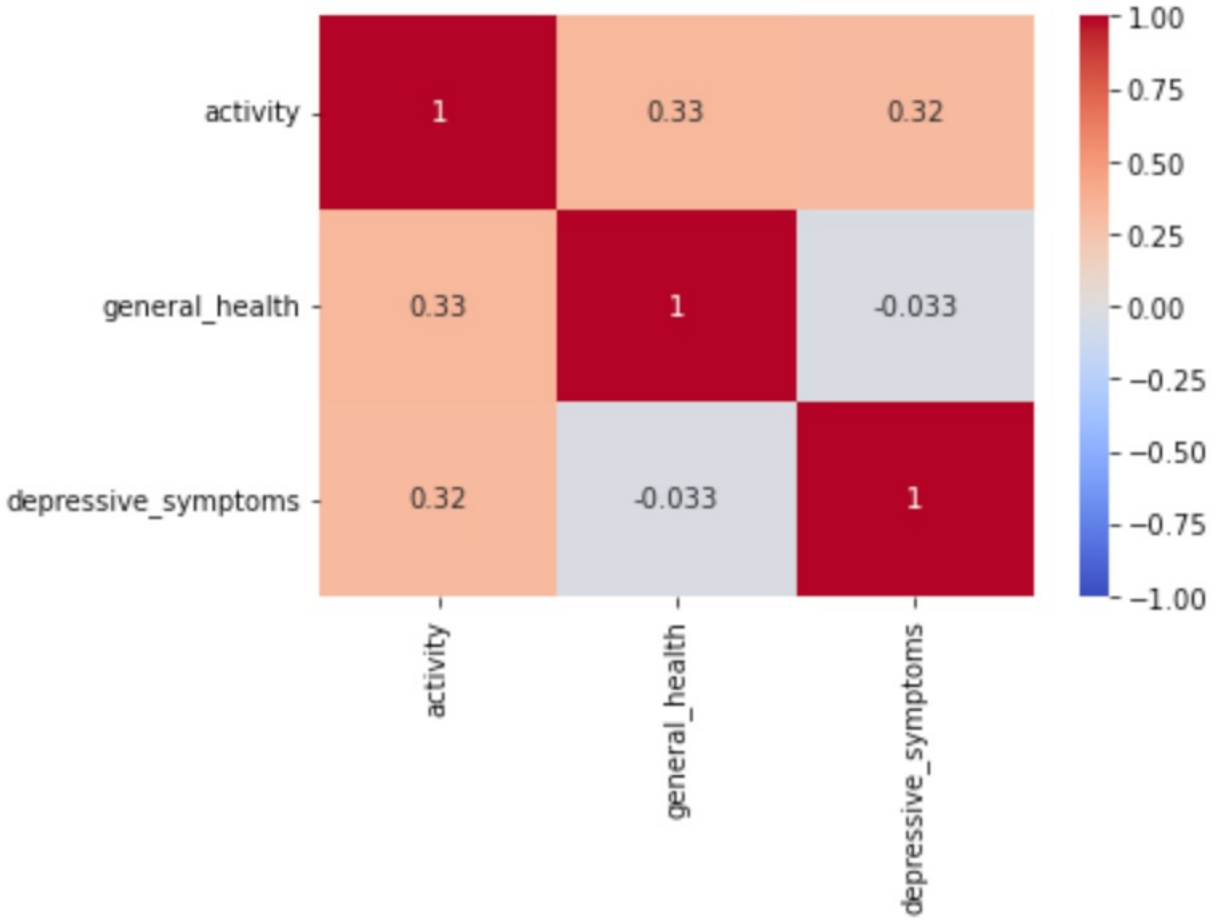






Results

Data Correlation



Discussion

Discussion

- Interesting insights
  - No significant trends from 2014 to 2019
  - Low variance in general health limits insights
  - Weak to moderate positive correlation between activity and general health
  - Weak to moderate positive correlation between activity and depressive symptoms
  - No significant correlation between general health and depressive symptoms
- Limitations
  - Low variance due to limiting research area
  - Other correlating factors (e.g. age, gender, education)
  - Self-report
  - Only depressive symptoms for mental health
  - Sample sizes, interval between measurements, generalization etc.

Friedrich-Alexander-Universität
Technische Fakultät

FAU

Physical Activity and Health in Europe

Leonie Färber
Methods of Advanced Data Engineering
WS 23/24