**Data Wrangling – Project Report**

**Impact of entertainment (Netflix and Spotify) on food consumption**

**Project group members**

1. Priya Jogie, pje230

2. [Angel Lokasa, VUnetID]

3. Jaspreet Singh, jsh480

4. Ikrame Zirar

**Research question**

[Formulate the question that you are addressing in your project here. Also mention the subquestions that you addressed along the way].

Hoeveel impact heeft entertainment consumptie (Netflix en Spotify) op voedselconsumptie?

How much impact does entertainment (Netflix and Spotify) consumption and happiness have on the amount of consumed food per country.

1. What is the most produced genre (Netflix) of a country and what is the relation with the value of the BMI per country?

2. What is the most produced genre of a country and what is the relation with the amount of food consumed per country?

3. Is there a correlation between the pace of the most listened to songs and the most produced genre per country.

4. Is there a correlation between what kind of genre and the amount of food consumed per country?

5. Is there a correlation between the amount of consumed food and the value of the BMI of a country.

* Per land welk genre (netflix) het meeste geproduceerd wordt en welk land het hoogste bmi heeft.

First we acquired the data frame of Netflix and we decided to create a new dataframe with the columns that we want to use.

We want to use the columns genre and the country. Because we want to merge these two and map them to other datasets. To do this efficiently we used the function copy().

* Per land welk genre (netflix) het meeste geproduceerd wordt en welk land het meeste voedsel consumeert
* Is er een correlatie tussen tempo van meest beluisterde liedjes per land en de genre (netflix) die het meest geproduceerd wordt in dat land
* is er een correlatie tussen wat voor muziek er geluisterd wordt (wat voor tempo, genre etc) en hoeveel voedsel er wordt geconsumeerd per land
* Correlatie tussen voedselconsumptie en bmi per land

**Data sources**

[Mention all the data sources that you used in addressing the research question, include the website, API and the last access data as well].

**Data wrangling methods**

[Describe the methods globally that you used to transform / wrangle the data in order to answer your research question. This could range from dealing with extracting the right information from the data format (e.g., JSON / XML), text processing (regular expressions), handling non-standard data (outliers, missing values), merging with other data (e.g., combining sales data to weather forecasts), special visualization (e.g., bokeh, D3, seaborn).

**Conclusion**

[State the conclusion of your research question. Also mention possible limitations in your approach that need to be addressed in future research.]