Market Research

La poule qui chante

Context

- La Poule Qui Chante – a French a French food company producing chicken products
- We want to develop internationally – but where?
- Large field of possibilities – this study will allow us to identify the best options.

Data for the study

- *Population*
- *Food availability*

Politics	Economics
Political stability	GDP per capita
Sociocultural context	**Technology**
Chicken available per capita	Agricultural credit
Ecology	**Legal**
Emissions from chicken production	Government investment in agriculture

https://www.fao.org/faostat/en/?#data

Preparation & Cleaning

General Steps

- Import the Python libraries
- Import the CSV files
- For each file:
  - Check the head and shape of the dataframe and then check the types of data
  - Check if there are null values
  - Delete the columns that will not be used for the analysis
  - Check if 'Zone'can be used as a primary key.
- Merge the files

Final merged file

- Find the null values:
  - GDP (1) – Taiwan data was missing. I used data from the IMF
  - Political stability (4) – I used IMF data for China and Taiwan. I used the political stability of France for New Caledonia and French Polynesia as they overseas territories of France.
  - Government investment (29) and agricultural credit (62) – the FAO explained that the response rates for these data are relatively low. I replaced the zero values with the average values of each column.
  - Emissions (4) - Djibouti and Maldives do not produce chicken(no emissions), FAO has data for Iceland for 2013 and 2019, and data for Latvia up to 2016. I have respectively interpolated and extrapolated the data.

Exploratory analysis of the final merged file

- Check type of data in each column
- Use ‘describe’ to check the les min, max, quartiles et standard deviation of each column
- Create a boxplot and a histogram for each column
- Check for outliers for each column

Results

Hierarchical Ascending Classification

- Data normalisation
- Create clusters using the Ward method
- Create a dendrogram with these clusters
- Determine the optimal number of clusters using the Davies-Bouldin index
- Create a contingency table and a heatmap
- Create a list of clusters with the countries in each cluster
- Perform deeper cluster analysis – “describe()” function, boxplots and histograms

Hierarchical Ascending Classification PESTEL analysis

Cluster 1

- Politics: Political stability - low
- Economics: GDP per capita - low
- Sociocultural context: Chicken available - low
- Technology: Agricultural credit - low
- Ecology: Emissions - high
- Legal: Government investment - low

Cluster 2

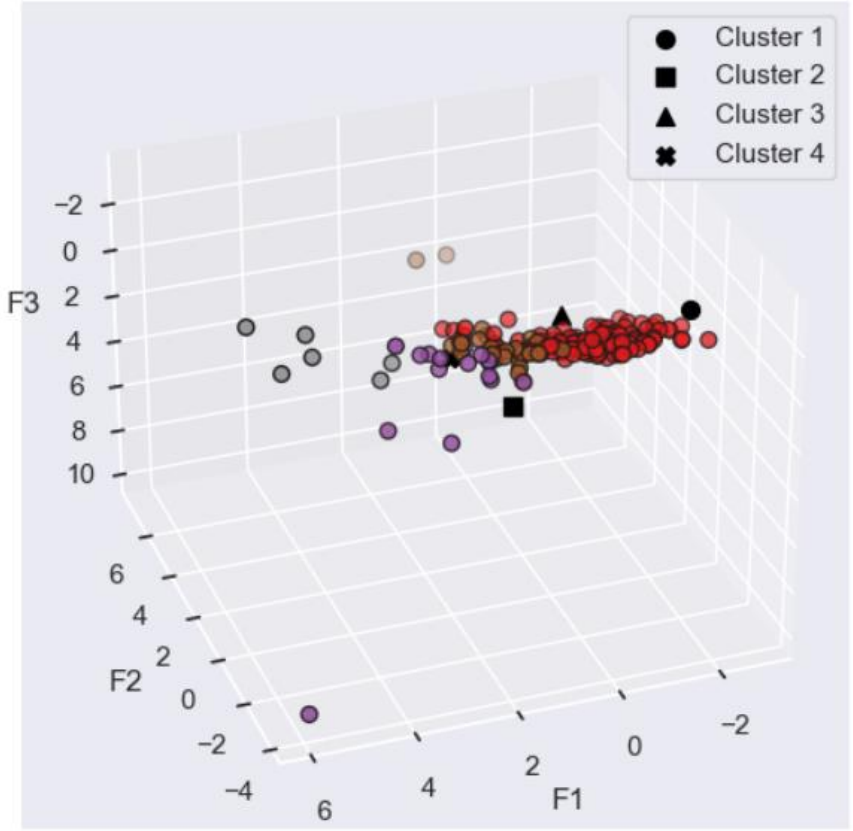
- Politics: Political stability - medium
- Economics: GDP per capita - medium
- Sociocultural context: Chicken available - high
- Technology: Agricultural credit - medium
- Ecology: Emissions - medium
- Legal: Government investment - medium

Cluster 3

- Politics: Political stability - high
- Economics: GDP per capita - high
- Sociocultural context: Chicken available - medium
- Technology: Agricultural credit - high
- Ecology: Emissions - low
- Legal: Government investment - High

K-means cluster analysis

- Determine the optimal number of clusters (4) – use the elbow method
- Use the optimal number of clusters to instantiate the k-means, then train them
- Add the centroids of each k-means cluster to a variable called “centroids”
- Scale column values and train PCA with scaled data and 4 components
- Visualize clusters and their centroids
- Create a list of clusters with the countries in each cluster
- Perform deeper cluster analysis – “describe()” function, boxplots and histograms



K-means cluster analysis

- Cluster centroid 1 is relatively far from cluster centroid 2 and cluster centroid 4.
- In contrast, the centroid of cluster 3 is close to 1, 2 and 4.
- The centroids of clusters 2 and 3 are the most similar clusters.

K-means cluster PESTEL analysis

Poorer Countries	High chicken consumption, high GDP	High chicken consumption, medium GDP	Richer countries
Politics: Political stability - low	Politics: Political stability - medium	Politics: Political stability - medium	Politics: Political stability - high
Economics: GDP per capita - low	Economics: GDP per capita - high (2nd)	Economics: GDP per capita - medium	Economics: GDP per capita – high (1st)
Sociocultural context: Chicken available - low	Sociocultural context: Chicken available - high	Sociocultural context: Chicken available - high	Sociocultural context: Chicken available - medium
Technology: Agricultural credit - low	Technology: Agricultural credit - high	Technology: Agricultural credit - medium	Technology: Agricultural credit - medium
Ecology: Emissions - high	Ecology: Emissions - low	Ecology: Emissions - medium	Ecology: Emissions - low
Legal: Government investment - low	Legal: Government investment - medium	Legal: Government investment - medium	Legal: Government investment - high

Countries to examine

HAC

- Finland
- Ireland
- Iceland
- Luxembourg
- Norway
- New Zealand
- Switzerland

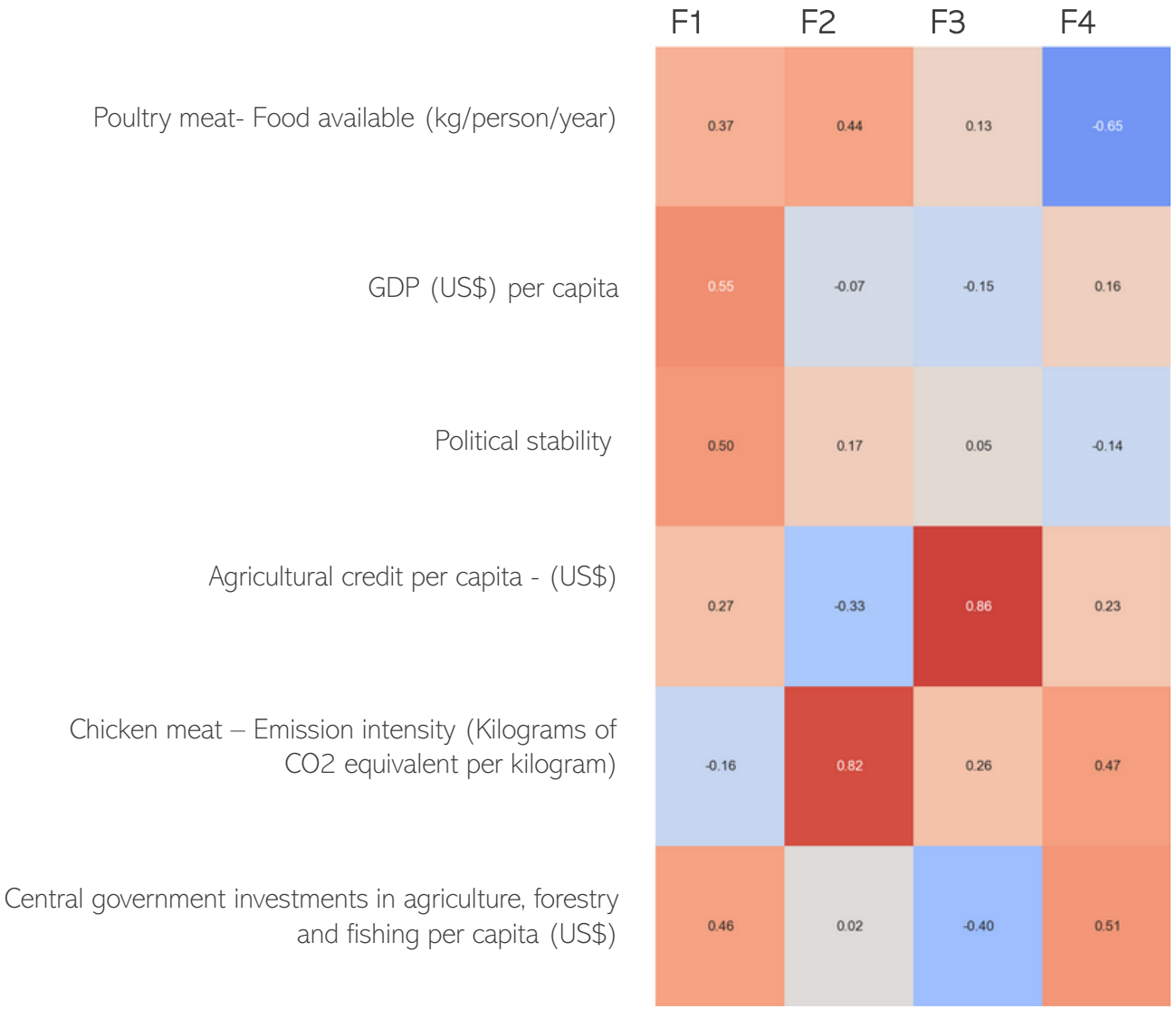
K-means

- Germany
- Australia
- Austria
- Belgium
- Belize
- Canada
- China - Hong Kong SAR
- Denmark
- Finland
- Israel
- Japan
- New Zealand
- The Netherlands
- United Kingdom
- Sweden
- United Arab Emirates
- United States of America

Principal Component Analysis

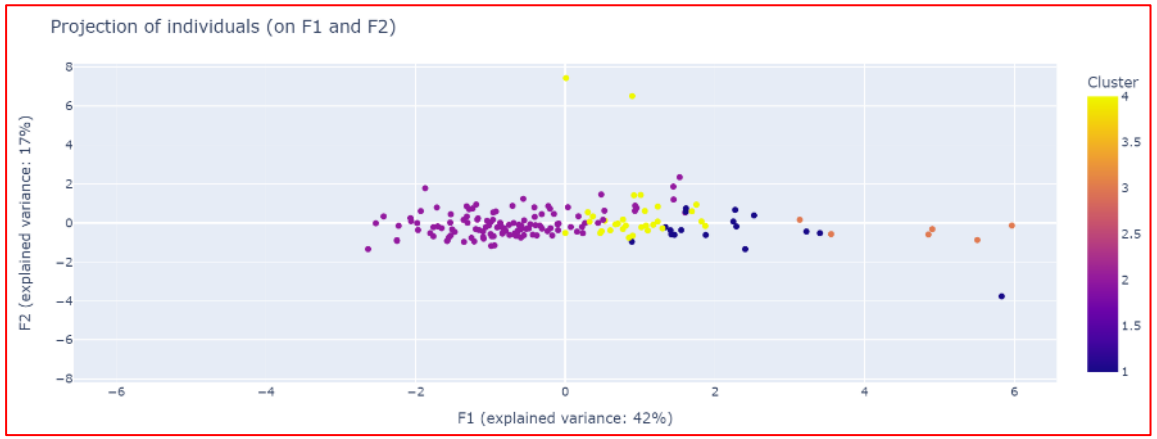
- Separate the quantitative values from the original dataframe.
- Train the scaler with these values and transform the data.
- Create a PCA and train it with the scaled data
- Calculate data variance to understand the contribution of each feature and determine the optimal number of components.
- Create a heatmap
- Draw a correlation circle and project the different countries onto the factorial planes.
- Analyse countries from these graphs.

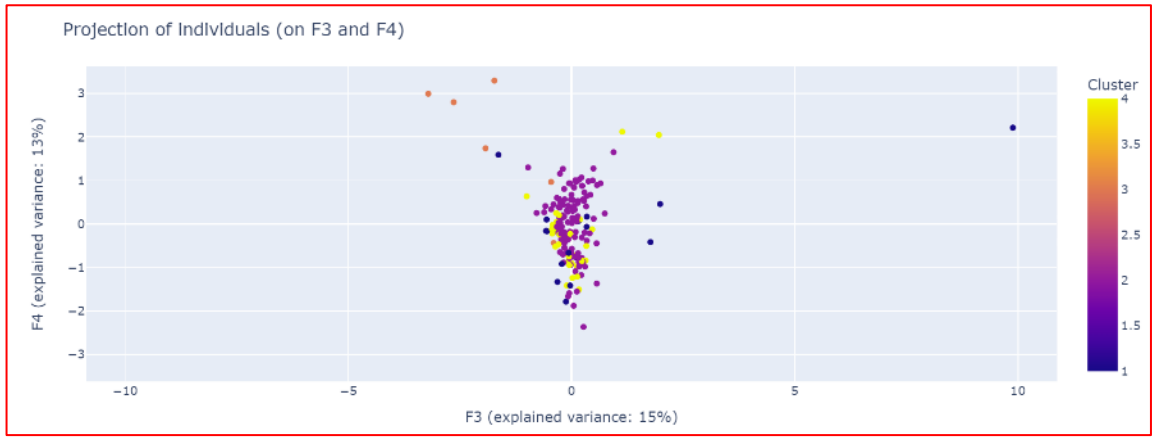
PCA – Principal Components



- F1 – Economic prosperity
- F2 – Quantity of emissions
- F3 – Agricultural credit
- F4 – Need for investment in agriculture

Component analysis – Visualisations – K-means

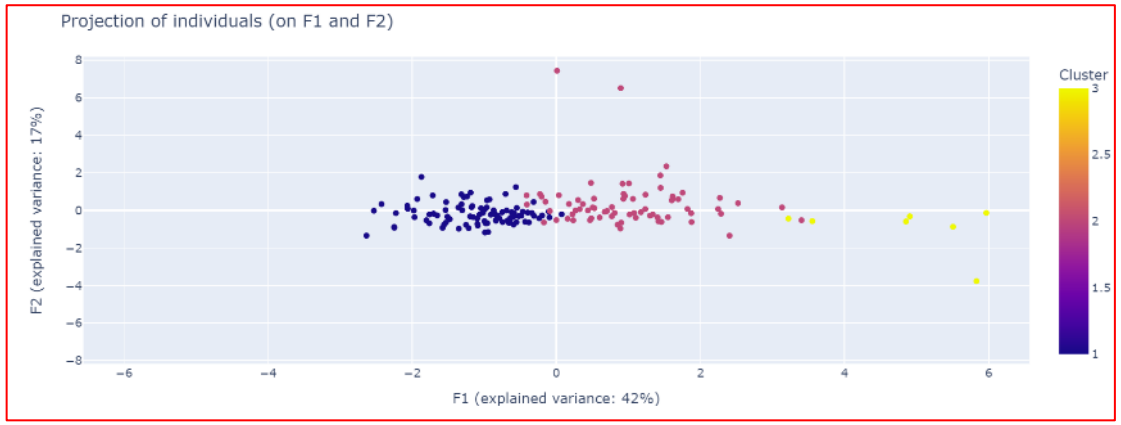


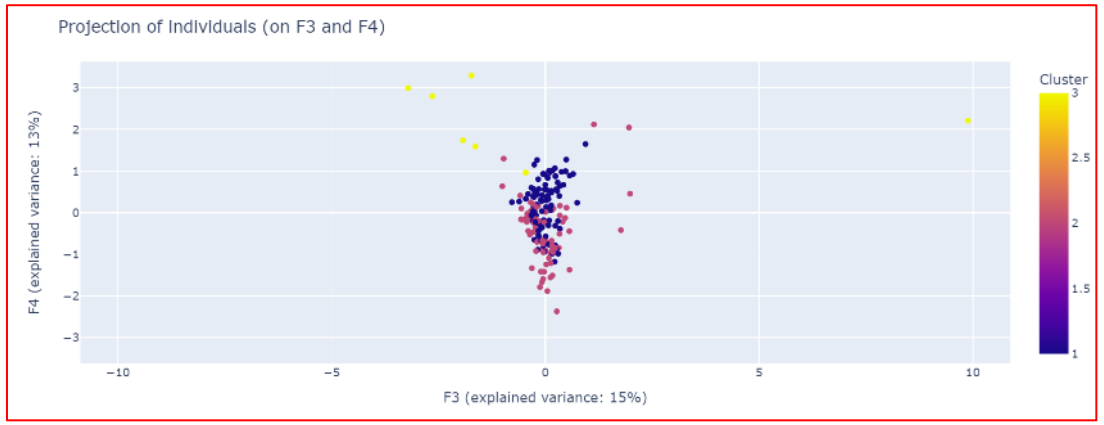


- F1 – Economic prosperity
- F2 – Quantity of emissions

- F3 – Agricultural credit
- F4 – Need for investment in agriculture

Component Analysis – Visualisations - HAC





- F1 – Economic prosperity
- F2 – Quantity of emissions

- F3 – Agricultural credit
- F4 – Need for investment in agriculture

International Development

Best regions

- Australasia (Australia and New Zealand)
  - A large population (24.5 million in Australia)
  - Lots of chicken available
  - Far from France – more complex regulations
- The European Union
  - Especially Netherlands and Ireland
  - Simpler and cheaper for product exports.
  - Less chicken available could suggest demand is lower than in Australasia