Project Report

# GitHub URL

(insert URL here)

# Abstract

(Short overview of the entire project and features)

This project explores Ireland's agricultural land management systems and the increase of "specialised dairy" farms, led by abolishing the European Union milk production quota in 2015. Objectives are to use Python libraries, such as NumPy and, Pandas, to identify the central farming regions in Ireland, to identify the extensively practised farm management systems, to estimate the change in agricultural land management followed by the abolishing of the EU milk production quota and to use Matplotlib for creating bar charts to visualise the results. Data used in the analysis are Teagasc National Farm Survey data for 2013 and 2016. Future research will use the regression analysis method to predict the probability of converting non-specialised dairy farms into "specialised dairying" and factors affecting the farmers’ choice of different land management systems.

# Introduction

(Explain why you chose this project use case)

Variables used in the analysis

The project analyses the distribution of “agricultural land management systems” across nine geographic regions in Ireland. This geographic regional categorisation is used by the Teagasc to identify the nature of agricultural activities in different parts of Ireland. The number of farms, the average size, predominant livestock and crop species, soil types and farming intensity (e.g. specialised or mixed species) varies across geographic regions.

The list of farm regions explored in the current project is depicted in table 1.

Table 1 List of Farm Regions in the National Farm Survey

|  |
| --- |
| West |
| Border |
| Mid-West |
| Midland |
| Dublin plus Mid East |
| South-West |
| South-East |
| Border, Midland and Western |
| Southern and Eastern |
| State |

This project analyses “agricultural land management systems” (also identified as "farm types" in this paper) currently practised in Ireland. Teagasc National Farm Survey data helps determine the number of farms and the average size of different “agricultural land management systems” across the geographic regions. Irish agricultural lands primarily consist of mixed-species management systems and specialised management systems, as listed in Table 2 below. Mixed-species management systems have a higher bio-diversity and are considered more resilient to climate change. They are more sustainable compared to specialised farms.

Table 2 List of Agricultural Land Management Systems (Farm Types) in National Farm Survey

|  |
| --- |
| Mixed crops and livestock |
| Mixed field crops |
| Mixed grazing livestock |
| Specialist tillage |
| Specialist dairying |
| Specialist sheep |
| Specialist beef production |
| Other |
| All farms |

Government policy

the European Union milk production quota capped Irish dairy production from 1984 to 2015 (Teagasc, 2015). After removing the milk production quota, the Irish dairy sector experienced a dramatic expansion. It is worthwhile to examine how "specialised dairying" increased, followed by the abolition of the EU milk production quota.

# Dataset

(Provide a description of your dataset and source. Also justify why you chose this source)

An analysis is conducted using secondary data collected by the National Farm Survey. This project employs three datasets which are available to the public via <data.gov.ie>, website managed by the Open Data Unit of the Department of Public Expenditure and Reform (DATA.GOV.IE, n.d.).

Selected datasets are appropriate because of the following benefits;

1. Readily available; no permission is required to access and use them
2. Three datasets are consistent and adopt the same data collection method and recording formats
3. Data is available for comparison for 2013 (i.e. before the EU milk production quota) and 2016 (i.e. after the EU milk production quota).

# Implementation Process

(Describe your entire process in detail)

1. Programming Environment: Firstly, a new Jupyter notebook was opened, then pandas and numpy library and the first dataset were imported to Jupyter notebook using pd.read\_csv() function. Dataset was imported from data.gov.ie website to the Jupyter notebook, and a name was assigned, i.e. ‘data1’ using the function below.

data1=pd.read\_csv("https://ws.cso.ie/public/api.restful/PxStat.Data.Cube\_API.ReadDataset/FSA01/CSV/1.0/en")

1. Column names were renamed to make them more understandable and appropriate using the function below.

data2=data1.rename(columns={"Farm Size":"FARM\_SIZE", "VALUE": "NUMBER\_OF\_FARMS\_AREA\_000"})

1. Data cleaning: checked the dataset for missing values and duplicates. There were no missing values or duplicates. Outliers were checked using histograms. Outliers are present in the dataset. They will be further explored in terms of the smallest and largest farms. The Mean value of the number of farms in Ireland is 14600. The following functions were used for cleaning the dataset;

Data3.isnull().sum()

data3.duplicated().sum()

data3.describe()

data3.hist()

1. Analysis: The first dataset contains values for 2013 and 2016. For the first part of the analysis, which focuses on identifying the major agricultural regions and management systems in Ireland, 2016 data was used. Firstly, DataFrame for all farms across different regions in Ireland in 2016 was created. The rows containing data for 2016 and rows containing “All Farms” were selected using the method explained in pandas (n.d.) and functions below;

data4=data3[data3["Year"] == 2016]

data5= data4[data4["FARM\_SIZE"] == "All farms"]

‘data5’ was sorted from smallest to largest, and the sorted DataFrame was assigned the name 'data6.'

data6 = data5.groupby("Region")["NUMBER\_OF\_FARMS\_AREA\_000"].sum().sort\_values()

Then the number of farms across different regions was compared using bar charts. data6.plot.bar() syntax; was used to visualise the region with the largest number of farms.

Matplotlib syntax;plt and the following functions were used to change the font size and to rotate the labels;

plt.rcParams.update({'font.size': 12})

plt.xticks(rotation=35, horizontalalignment="right")

Matplotlib syntax;plt and the following functions were used to add the chart title and axis labels.

plt.title("Distribution of Farms in Ireland in 2016")

plt.xlabel("Farming Region")

plt.ylabel("Number of Farms (000)")

Next, (FARM\_SIZE) column in the dataset and .plot () syntax were used for visualising the distribution of the largest and smallest farms in Ireland.

The distribution of large farms (i.e. farms having at least 100Ha land area) was plotted against different regions using the function below;

data8.plot(kind='bar', x='Region', y='NUMBER\_OF\_FARMS\_AREA\_000', color="brown")

The distribution of small farms (i.e. farms with less than 10Ha of land area) was plotted against different regions- using the function below;

data10.plot(kind='bar', x='Region', y='NUMBER\_OF\_FARMS\_AREA\_000', color="yellow")

The second dataset contains the number of farms across “agricultural land management systems” and regions, and it is downloaded and stored on the local device. It was imported to Jupyter Notebook and assigned a name ‘Facts1’using the following function;

Facts1=pd.read\_excel(r"C:\Users\Chaagi\Desktop\PhD UCD\D\Data Analytics\NFS02.xlsx")

Data cleaning was conducted on the new dataset using the previous steps. Then the second data set was used to find the predominant type of “agricultural land management systems” practised in the largest farming region in Ireland, i.e. Border, Midland and Western. The largest farming region in Ireland was identified in the previous analysis. The “agricultural land management systems” data for 2016 in the Border, Midland and Western regions was retrieved from the dataset and plotted using the same steps and functions described previously.

Furthermore, the most extensively practised farm management system in Ireland was also identified using this second dataset and the function below;

Total\_Land\_2016\_Sort.plot(kind='barh', x='FARM\_TYPE', y='LAND\_SIZE\_000', color="red")

The third dataset contains the average size of farms and “farm types”. The second dataset contains the number of farms and “farm types”. Therefore, the second and third datasets have a common column ('FARM\_TYPE'), so they can be merged. The second and Third datasets were merged using pd.merge() function, and the resulting table contains the farm types, number of farms and average size. The Merged DataFrame was assigned the name 'Tab' using the function below;

Tab=pd.merge(Facts10, Figures6, on="FARM\_TYPE", how="inner")

In the merged DataFrame, the column containing the number of farms was multiplied by the column containing average size to estimate the total land area under each farm type. And assigned the name ‘Tot’

Tot=Tab["NUMBER\_OF\_FARMS\_TYPE\_000"] \* Tab["AVERAGE\_FARM\_AREA\_000"]

‘Total\_Land\_2016’ was assigned to the DataFrame containing Total land areas across farm types.

This process was repeated for year 2013 (resulting variable named ‘Total\_Land\_2013’) and 2016 (resulting variable named ‘Total\_Land\_2016’).

An Unstacked bar plot showing the total agricultural land areas in 2013 and 2016 was created. A new variable was defined as 'Total\_Land\_2013\_2016', which was used to visualise the land areas across farm types in 2013 and 2016. The function used for the unstacked bar plot was developed based on Shane Lynn (n.d.) method and is given below;

Total\_Land\_2013\_2016 = pd.DataFrame({'Total\_Land\_2013': [4537.00, 301.04, 872.04, 2127.60, 444.00, 482.08, 152.62, 215.67, 20.80],

'Total\_Land\_2016': [4455.00, 275.42, 953.12, 2074.95, 437.90, 426.88, 123.27, 126.28, 33.02]}, index=['All farms', 'Specialist tillage', 'Specialist dairying', 'Specialist beef production', 'Specialist sheep', 'Mixed grazing livestock', 'Mixed crops and livestock', 'Mixed field crops', 'Other'])

# Results

(Include the charts and describe them)

Figure 1 shows the number of farms in different regions across Ireland. Accordingly, in 2016, the highest number of farms are found in the "Border, Midland and Western" regions (72500 farms), followed by Southern and Eastern (65,000 farms) and West (31400 farms). The lowest number of farms are found in the Dublin plus Mid East region (10300 farms).

# 

Figure 1 Number of Farms Across Agricultural Regions in Ireland in 2016

# Figure 2 shows the number of farms in Ireland that are larger than or equal to 100ha (>=100ha). Accordingly, The large farms are most frequently found in the Southern and Eastern region (3400 farms), followed by Border, Midland and Western part (1400 farms) and South-east (1100 farms) in descending order. The West region has the lowest number of large farms (400 farms).

# A picture containing icon Description automatically generated

Figure 2 Distribution of Large Farms (=>100Ha) Across Agricultural Regions in Ireland

# Figure 3 shows the distribution of small farms (i.e. farms at least 10Ha in size) across the regions. Accordingly, most small farms are located in the Border, Midland and Western part.

# Shape, rectangle Description automatically generated with medium confidence

Figure 3 Distribution of Small Farms (=>10Ha) in Ireland in 2016

# Figure 4 shows the types of farms available in the Border, Midland and Western region. Accordingly, the most common type of farming in the Border, Midland and Western regions is "Specialist beef" production (46000 farms), followed by "Specialist sheep" (10300 farms) and "Mixed grazing livestock" (6500).

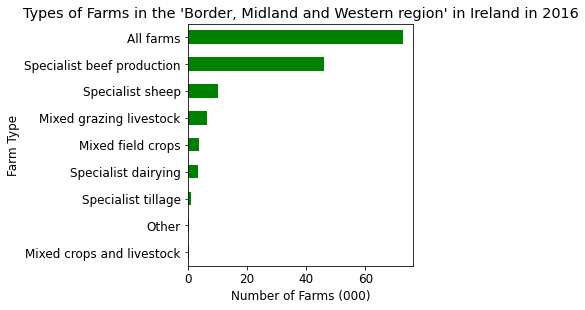


Figure 4 Types of Farms found in the Main Farming region in Ireland

Figure 5 shows Ireland's total agricultural land area in 2016. Accordingly, most of the agricultural lands in Ireland are "Specialist beef farms" (2074.95 Ha,000) and "Specialist dairying farms" (953.12 Ha,000). The total land area under Mixed type land management systems (i.e. Mixed crops and livestock, Mixed field crops, Mixed grazing livestock) is 15.2% of all agricultural lands in Ireland.

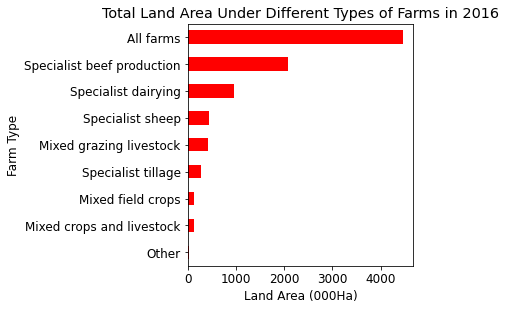


Figure 5 Total land area managed under different agricultural systems.

# Figure 6 shows the area under different “agricultural land management systems” in 2013 and 2016. Accordingly, the total agricultural land area slightly decreased in 2016 compared to 2013. Land areas under all farming types have declined except for "specialist dairying", which shows an increment. Compared to 2013, the increase in specialised dairying was 9.3% or 81080Ha in 2016.

# 

Figure 6 Land Area Under Agriculture in 2013 and 2016

# Insights

(Point out at least 5 insights in bullet points)

* The highest number of farms are located in Ireland's Border, Midland and Western regions. Most farms (63.5%) in this region are specialised in beef production.
* A majority (60.3%) of the small farms (small farms <=10Ha) are also located in the Border, Midland and Western region.
* The majority (69.4%) of the large farms (large farms are =>100 ha) are located in the Southern and Eastern regions, where the second largest number of farms are located.
* The total land area under farming slightly declined (1.8% decline on 2013 value) in 2016. The decline is evident in both mixed-species management systems and specialised management systems, except for specialised dairying.
* The total farm area managed as specialised dairy increased by 9.3% post-EU Milk quota in 2016 compared to during the implementation of the EU Milk quota (2013).

# References

(Include any references if required)

DATA.GOV.IE. (n.d.) *Datasets.* Available at: https://data.gov.ie/dataset/ (Accessed 26 Sep. 2022).

pandas (n.d.). *How do I select a subset of a DataFrame?* Available at: <https://pandas.pydata.org/docs/getting_started/intro_tutorials/> [Accessed 26 Sep. 2022].

Shane Lynn(n.d.) *Make Better Bar Charts in Python using Pandas Plot.* Available at: https://www.shanelynn.ie/bar-plots-in-python-using-pandas-dataframes/ (Accessed 26 Sep. 2022).

‌Teagasc (2015) *‘The End of the Quota Era: A History of the Irish Dairy Sector and Its Future Prospects*. Athenry: Agricultural Economics and Farm Surveys Department. Available at: <https://www.teagasc.ie/media/website/publications/2015/> [Accessed 26 Sep. 2022].