**Health graduate outcomes analysis**

**Introduction**

The Health Graduate Outcomes dataset provides valuable insights into the employment and educational status of health graduates over time. Understanding these outcomes is crucial for policymakers, educational institutions, and stakeholders to assess the effectiveness of health education programs and to identify areas for improvement. This report aims to analyse the dataset to uncover significant trends, relationships, and patterns that can inform future decisions, by focusing on the following objectives:

1. Trends in graduate outcomes over time.
2. Gender differences in graduate outcomes.
3. Impact of years since graduation on outcomes.
4. Three statistical tests of field specific differences in graduate outcomes.

The findings will help provide insights into the performance of health graduates and may assist educational institutions and policymakers in improving graduate support and career services.

**Dataset Overview**

The dataset under analysis pertains to health graduate outcomes in Ireland, focusing on various dimensions such as gender, field of education, and years since graduation. This dataset is publicly available through the Central Statistics Office (CSO) of Ireland and is accessible under the Creative Commons Attribution 4.0 license

The data has been compiled by the CSO, utilising administrative records to track the post-graduation activities of health graduates. This includes employment status, further education enrolment, and periods where graduates are not captured in Irish administrative data, which may indicate emigration. The dataset encompasses graduates from various health-related fields, including medicine, nursing, and midwifery, over multiple graduation years. The dataset is available for public use and can be accessed through Ireland's open data portal Data.gov.ie.

The dataset contain 4,356 records and 14 columns, detailing various aspects of health graduates post-education experiences. The main 8 columns that I used for my analyses are:

1. STATISTIC: Type of statistic reported.
2. Statistic Label: Description of the statistic.
3. Graduation Year: Year of graduation.
4. Fields of Education and Training: Specific field of study.
5. Years since Graduation: Time elapsed since graduation.
6. Graduate Outcomes: Employment and education status.
7. Gender: Gender of the graduate.
8. VALUE: Number of graduates.

The dataset includes categorical variables such as Fields of Education and Training, Graduate Outcomes, and Gender, as well as numerical variables like Graduation Year, Years since Graduation,

**1. Data Exploration and Cleaning**

Initial exploration revealed that the dataset contains no missing values. The Graduate Outcomes column includes categories such as Employment only, Employment and Education, Education only, Neither Employment nor Education, Not Captured, and All Graduate Outcomes. The VALUE column represents the number of graduates in each category.

**2. Descriptive Statistics**

The VALUE column has mean of 106.48, median of 0, max of 1,470, and min of 0. These statistics suggest a highly skewed distribution, with a significant number of records reporting zero graduates.

A screenshot of a computer

AI-generated content may be incorrect.

**Data Analysis and Insights**

**Insight 1: Trends in Graduate Outcomes Over Time**

Grouping the data by Graduation Year and Graduate Outcomes and summing the VALUE column reveals trends in graduate outcomes over time. A bar plot of this aggregated data illustrates the distribution of graduates across different outcomes for each graduation year.

A graph with different colored lines

Description automatically generated  
This chart shows graduate outcomes from 2010 to 2020. The blue bars All Graduate Outcomes are the largest each year, showing the total graduates tracked.Most graduates are in the Employment only in red, fewer graduates combine Employment and Education in green, Education only in orange is the smallest group and the Not Captured category in brown is consistently low.There is a noticeable drop in graduate numbers after 2017, possibly due to fewer graduates or changes in data collection.

**Insight 2: Gender Differences in Graduate Outcomes**

Aggregating the data by Gender and Graduate Outcomes and summing the VALUE column provides insights into how graduate outcomes change between genders. A bar plot of this data highlights the number of graduates in each outcome category for males and females.

**A graph of a graduate

Description automatically generated with medium confidence**

This bar chart compares the number of graduates in different outcomes like employment, education, or neither based on gender. It shows all genders combined have the highest overall numbers for all outcomes. Females have more graduates in most categories compared to males, males have fewer graduates overall but a similar trend of outcomes. Gender seems to impact graduate outcomes, with females having higher numbers in some categories.

**Insight 3: Impact of Years Since Graduation on Outcomes**

Grouping the data by Years since Graduation and Graduate Outcomes, by summing the VALUE column offers insights into how graduate outcomes evolve over time. A line plot of this data shows the number of graduates in each outcome category across different years since graduation.

**A graph of a graduate

Description automatically generated with medium confidence**

This line chart tracks how graduate outcomes change over time since graduation. It shows, the total number of graduates in employment or education decreases steadily over time. The category Not Captured graduates not tracked stays relatively low. Neither employment nor education remains flat, showing minimal change. Over time, fewer graduates remain in the employment or education categories, indicating long-term shifts.

**Insight 4: Three Statistical Tests of Field Specific Differences in Graduate Outcomes**

**ANOVA Test**

To assess the significance of differences in graduate outcomes across different fields of education, a one-way ANOVA was conducted. The results show a large difference in the average graduate outcomes between different fields of education.

**ANOVA Result:** F-statistic = 196.82910517827392 p-value = 9.105770819656876e-44

There is a big difference in the average graduate outcomes between different fields of education.

**A graph showing a number of classes

Description automatically generated with medium confidence**

The distribution shows a lower range of graduate outcomes with a tighter spread, indicating consistency in graduate values. There are a few outliers present but at relatively lower levels compared to Nursing and Midwifery. In other hand Nursing and Midwifery has a wider spread in the data, with a larger number of high value outliers. The central data range indicates a generally higher median compared to Medicine, reflecting greater variability in outcomes.

**T-Test**

A t-test was performed to compare graduate outcomes between genders. The results showed a large difference in the graduate outcomes between males and females.

**T-test Result:** t-statistic = -13.67506350274108 p-value = 2.6539693791275286e-41

There is a big difference in the graduate outcomes between males and females.

A graph of a graph showing genders

Description automatically generated with medium confidence

The data for male graduates shows a narrow distribution, with lower median values and a few outliers. This suggests relatively consistent graduate outcomes within this group. Female graduates exhibit a wider distribution with a higher median compared to males. There are more extreme outliers, indicating greater variability in outcomes among females. The combines data from all genders, showing the broadest range and a slightly higher median than the individual gender categories. The presence of numerous high value outliers suggests that certain graduate outcomes skew the data.

**Chi-Square**

A chi-square test was also managed to check the relationship between gender and graduate outcomes. The results indicated no big relationship between these variables.

**Code Example:**

**Chi-Square Result:** chi2-statistic = 0.0 p-value = 1.0

There is no big relationship between Graduate Outcomes and Gender.

**A graph of a graduate

Description automatically generated**

This scatterplot shows how median graduate outcomes change over years since graduation, categorised by outcomes like Employment and Education, Education Only, and others. Outcomes are more varied in the first few years 1-2 years after graduation. Median values generally decrease over time. Categories like Not Captured and Neither Employment nor Education have consistently lower outcomes.

**Challenges and Solutions**

1. During the analysis, several challenges were faced are skewed distribution of VALUE, the VALUE column exhibited a highly skewed distribution, with many records reporting zero graduates. This was addressed by focusing on aggregated data and visualisations to identify meaningful patterns.
2. One of the challenges I faced during this analysis was understanding some of the columns in the dataset, such as **C03143V03795**, **C03433V04141**, **C03544V04280**,and **C03919V04671**. These column names were not clear, and there was no documentation to explain what they represented, without knowing their purpose or meaning I was unable to analyse or include them in the report. To ensure accuracy, I decided to focus on columns that I could clearly understand and interpret. However, this limited the scope of my analysis and was a learning point for me about the importance of having well-documented datasets.
3. Handling of Not Captured Data, the Not Captured category in Graduate Outcomes was considered as missing data and excluded from certain analyses to maintain the integrity of the results.

**Conclusion**

This analysis of the Health Graduate Outcomes dataset provides valuable insights into the career paths of health graduates in Ireland. Understanding these trends is crucial for educational institutions and policymakers to develop strategies that support graduates in their transition from education to employment or further study.