

CITS3401 Data Warehousing

Project 1

Dataset: US Adult Income

Dimensions:

- Age
- Gender
- Race
- Native Country
- Education
- Occupation
- Marital Status
- Income Bracket

Measures:

- Number of people (count) - used primarily instead of Fnlwgt in this report
- Fnlwgt (people represented by entry)
- Hours worked Per Week
- Capital Gain
- Capital Loss

Business Queries

Query 1

Show the percentage of adults in each Income Bracket, grouped by their Marriage Category (Married and Unmarried)

Query 2

How many adults aged 25 and under earn more than \$50,000? Show the number per gender

Query 3

Show the number of adults in each Income Bracket, grouped by:

- Age Bracket
- Educational Institution they attended
- Occupation Type

Query 4

What are the Average Number of Hours worked per Week by Occupation Type and Gender?

Query 5

What is the average Capital Gain and Capital Loss according to:

- Age Bracket
- Education Level
- Occupation Type

Query 6

Show the percentage of Adults per Educational Institution, categorized by whether their Native Country is classified as Developing or Developed

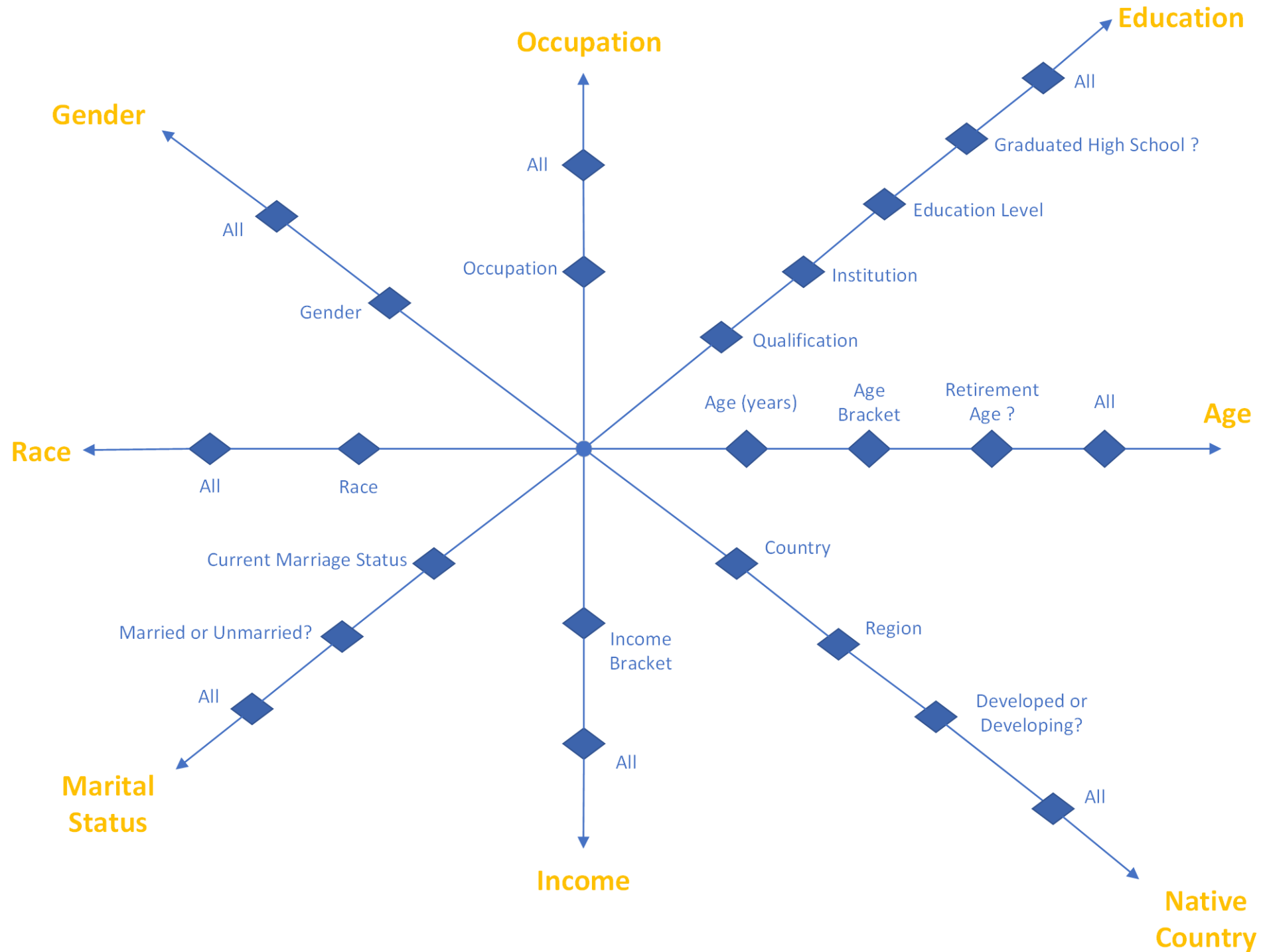
Query 7

Show the Education classification for Adults from each region

Query 8

Show the percentage of adults per occupation type, categorized by whether they graduated from High School or not.
Show which income bracket they are in

StarNet Diagram

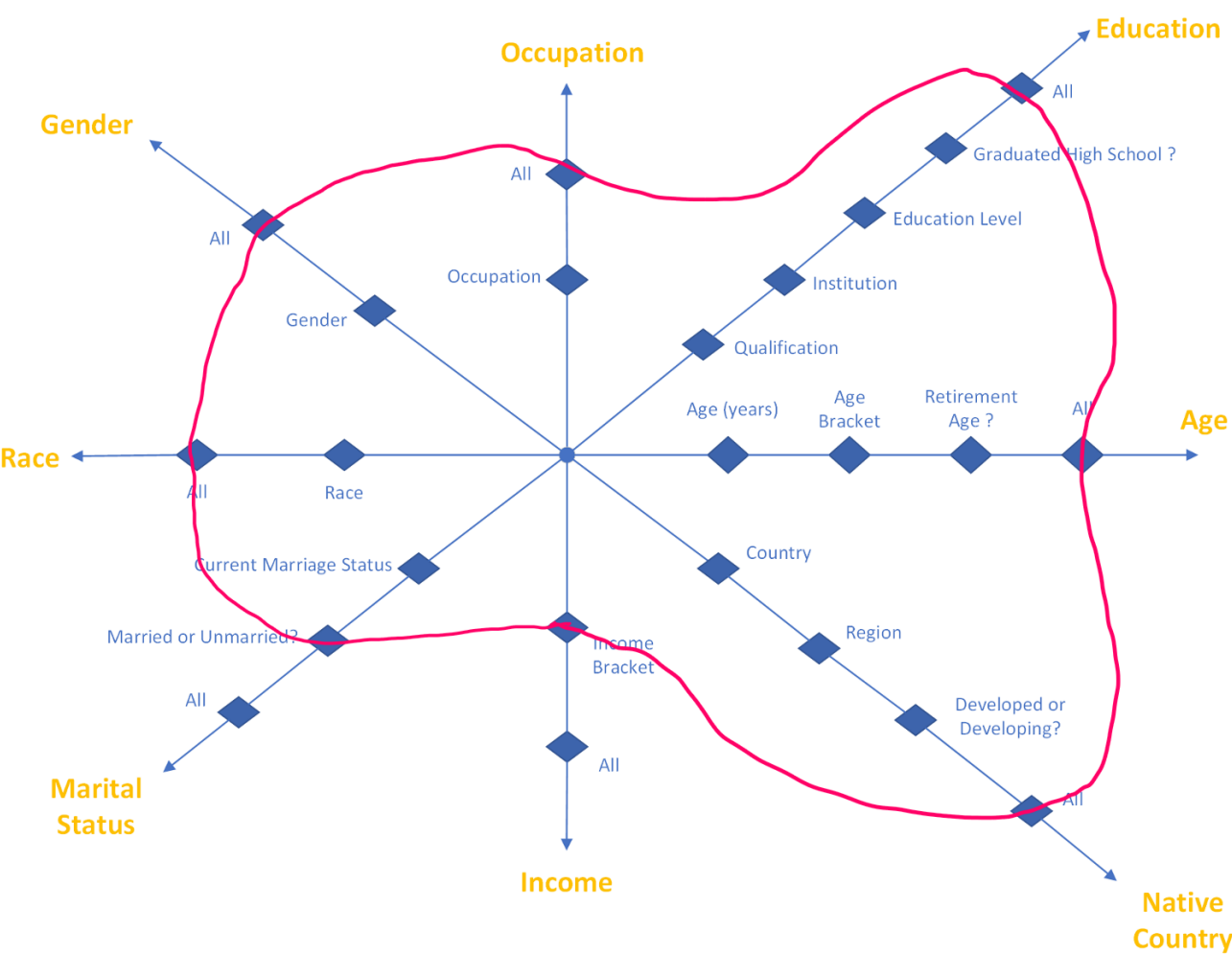


StarNet Footprints

Business Queries 1 and 2

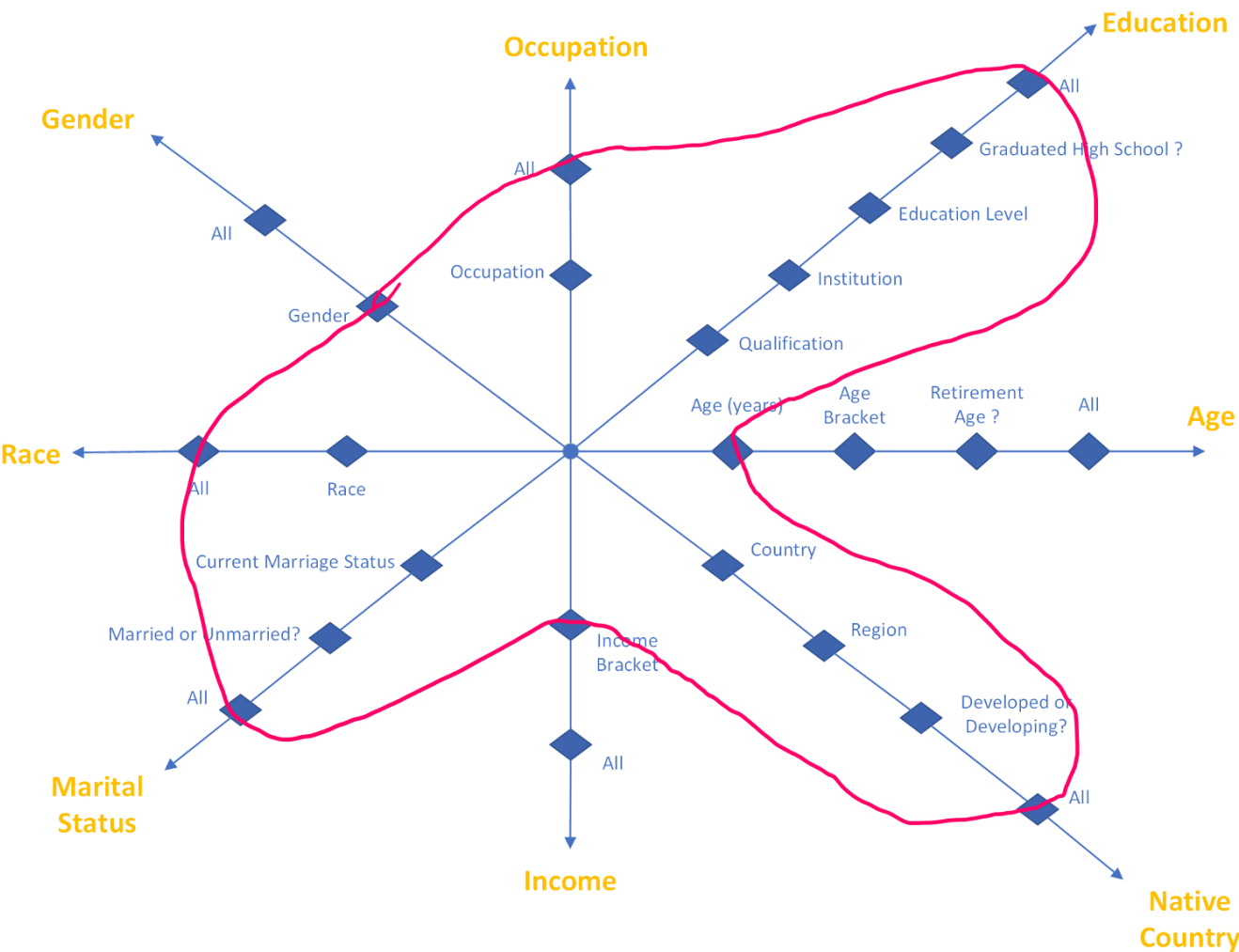
Business Query 1

Show the percentage of adults in each Income Bracket, grouped by their Marriage Category (Married and Unmarried)



Business Query 2

How many adults aged 25 and under earn more than \$50,000?
Show the number per gender



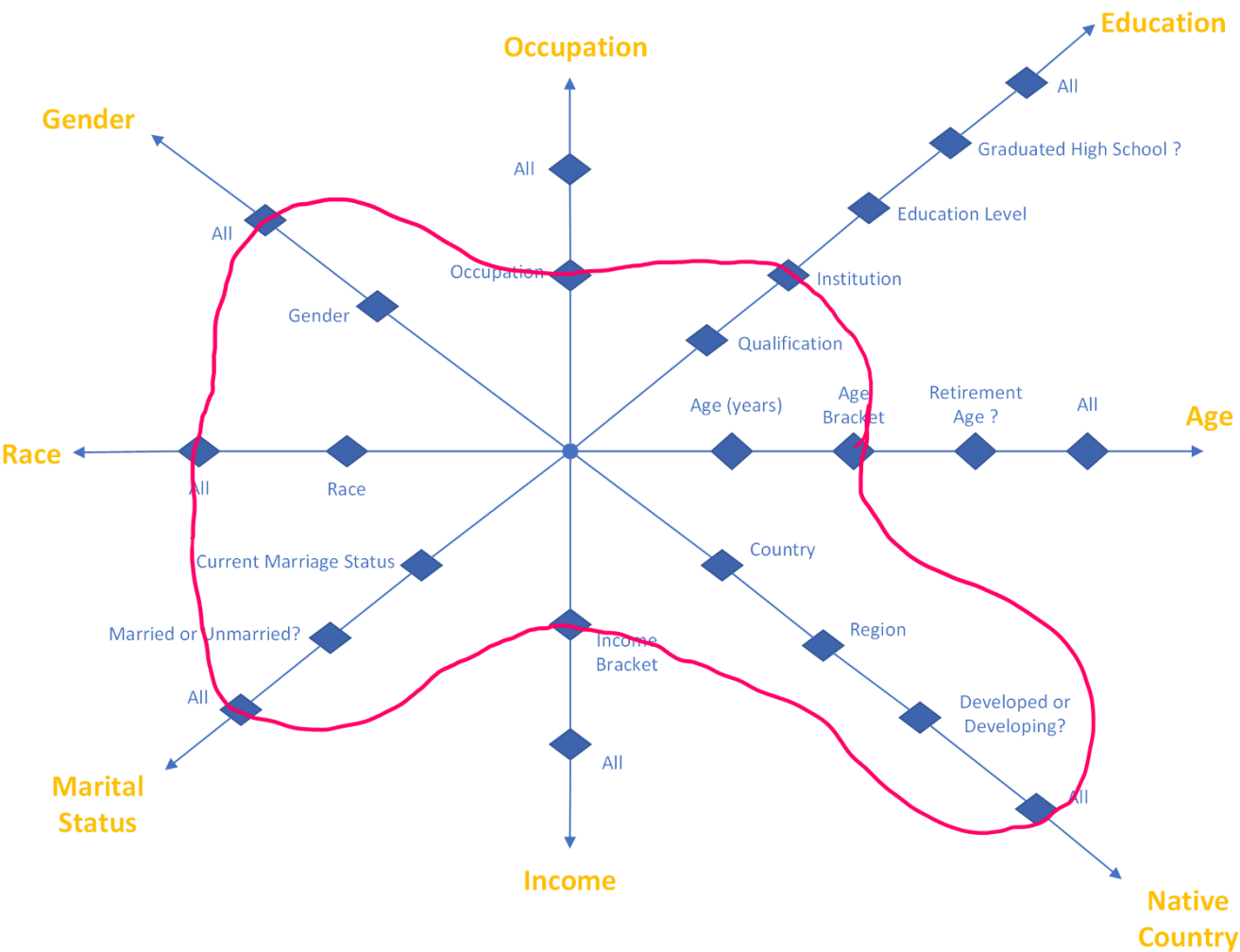
StarNet Footprints

Business Queries 3 and 4

Business Query 3

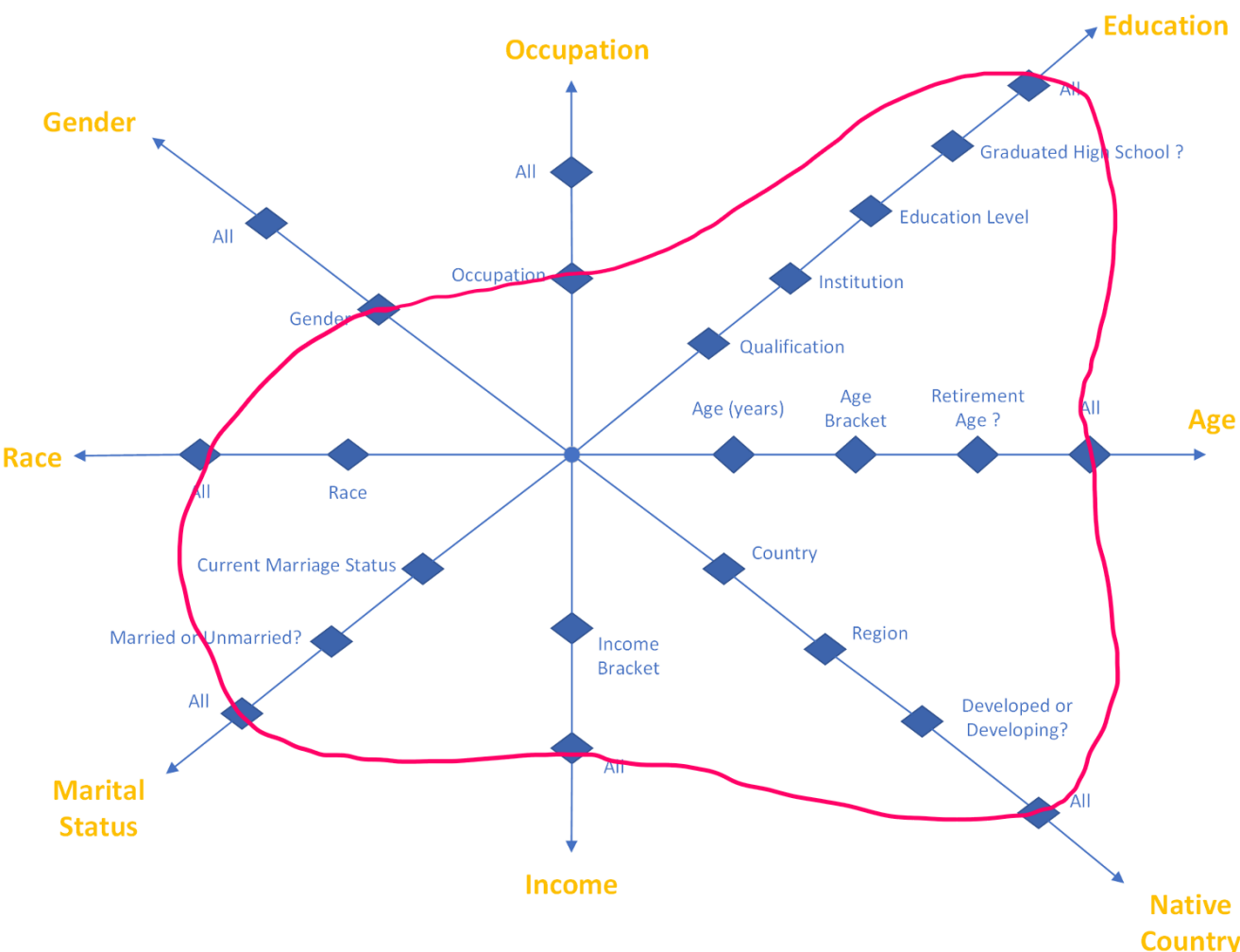
Show the number of adults in each Income Bracket, grouped by:

- Age Bracket
- Educational Institution they attended
- Occupation Type



Business Query 4

What are the Average Number of Hours worked per Week by Occupation Type and Gender?



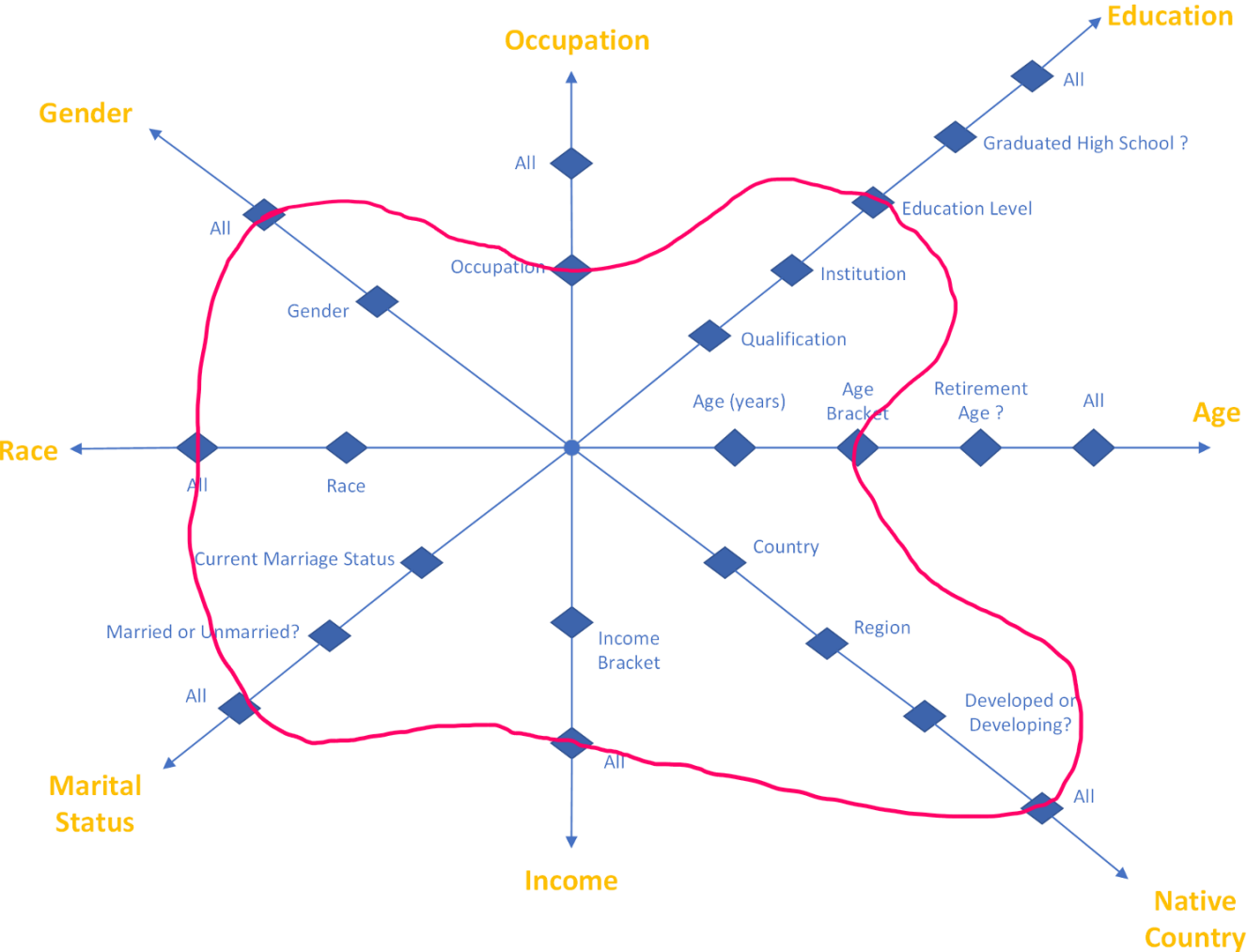
StarNet Footprints

Business Queries 5 and 6

Business Query 5

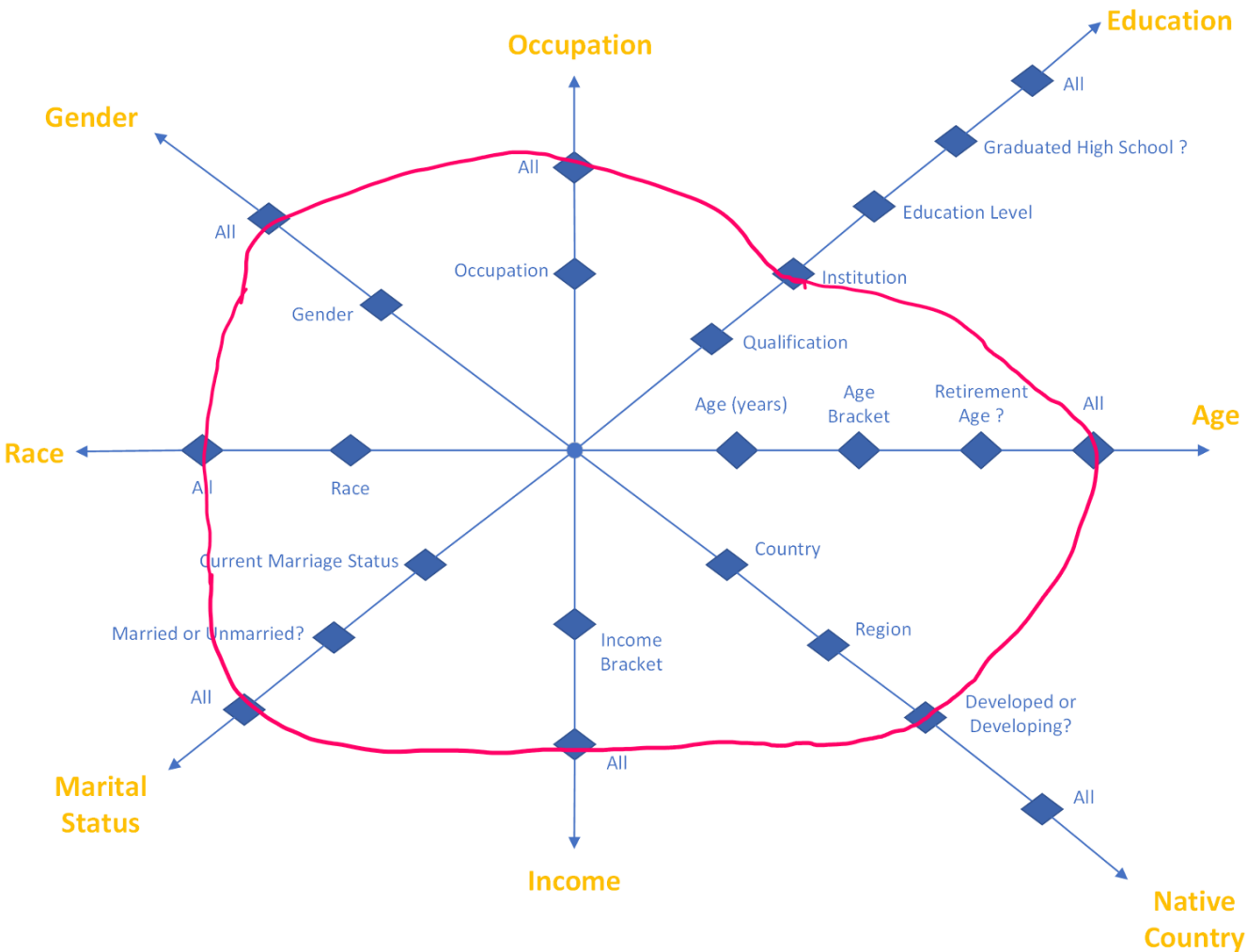
What is the average Capital Gain and Capital Loss according to:

- Age Bracket
- Education Level
- Occupation Type



Business Query 6

Show the percentage of Adults per Educational Institution, categorized by whether their Native Country is classified as Developing or Developed

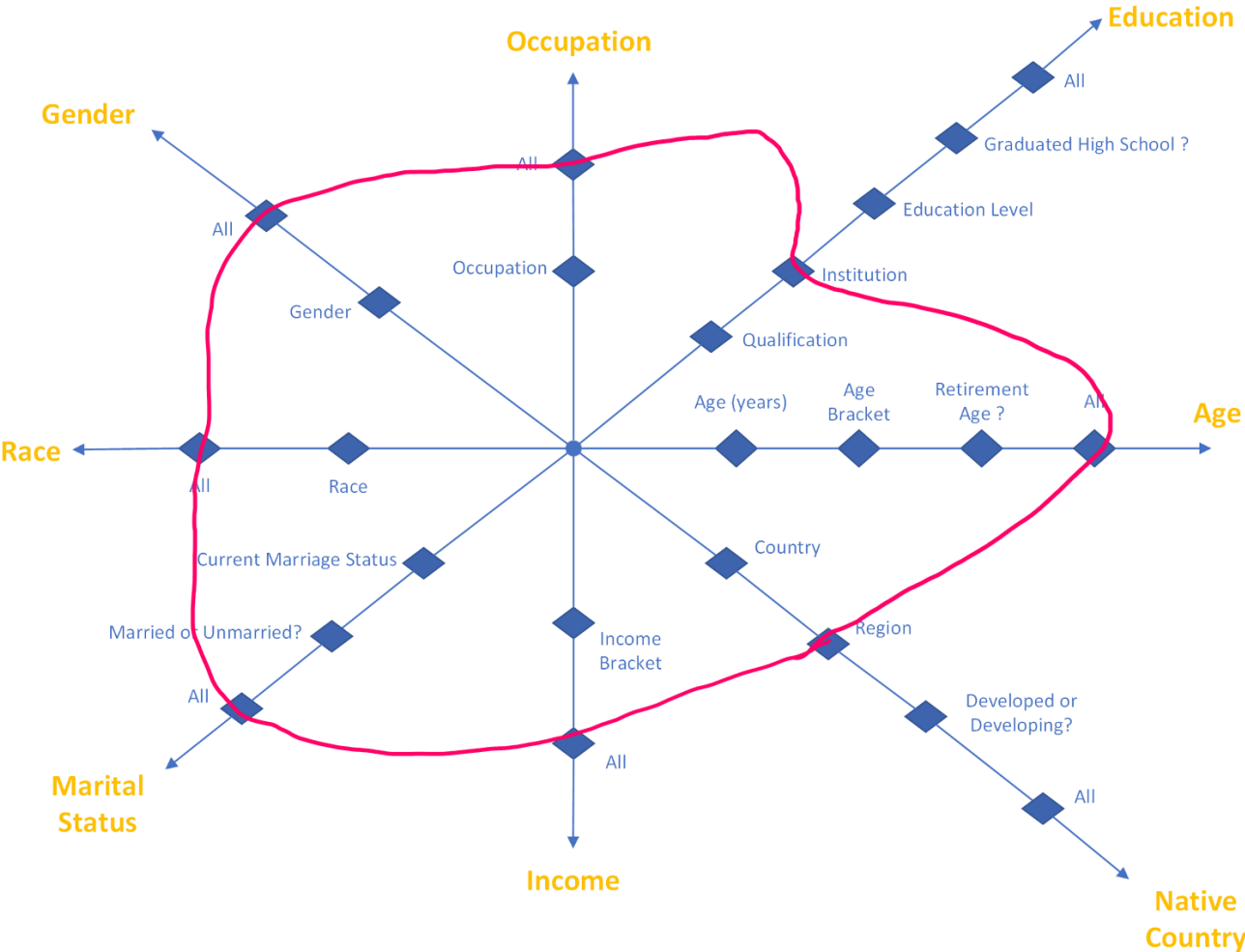


StarNet Footprints

Business Queries 7 and 8

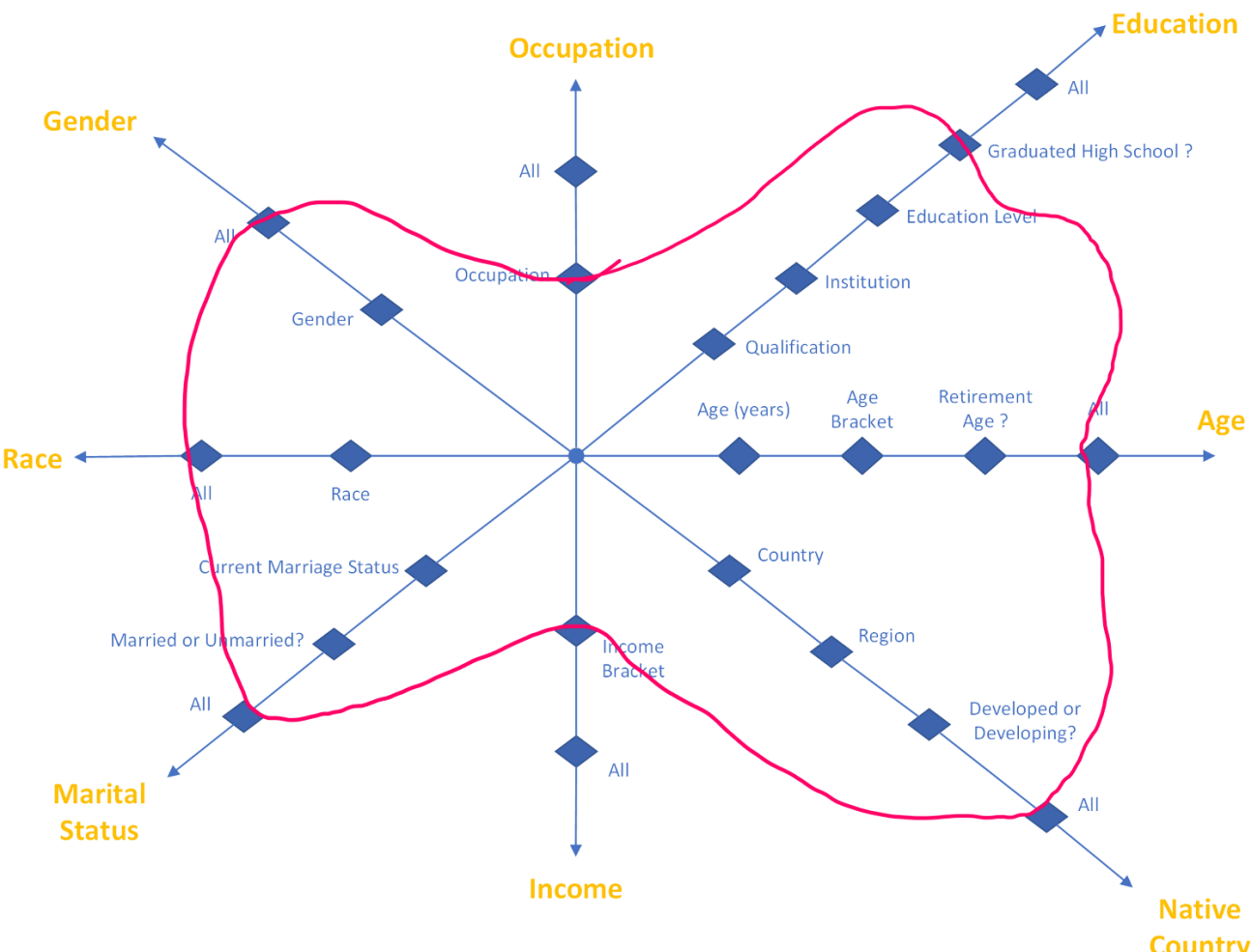
Business Query 7

Show the Education classification for Adults from each region



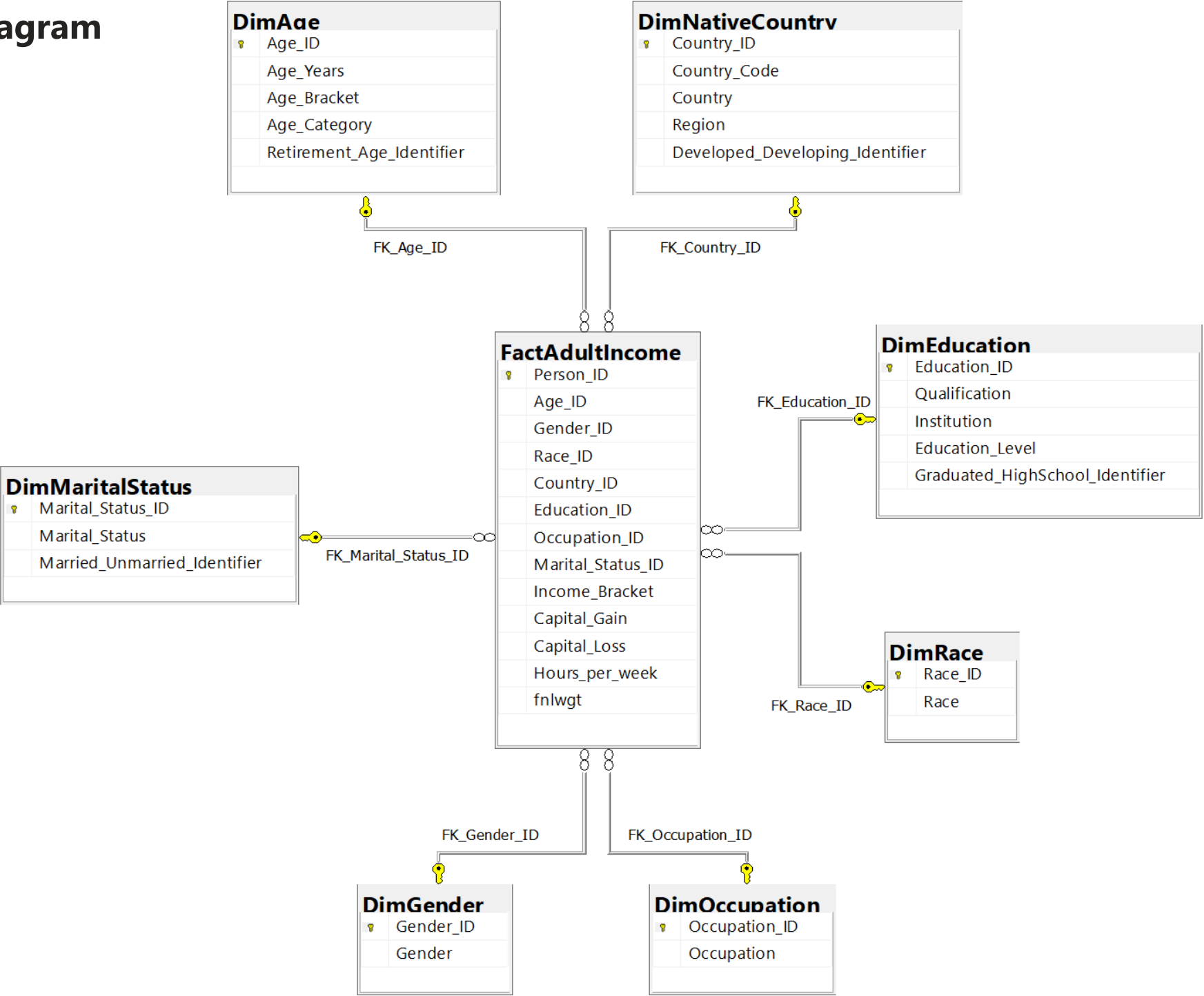
Business Query 8

Show the percentage of adults per occupation type, categorized by whether they graduated from High School or not.
Show which income bracket they are in



Data Warehouse Diagram

Star Schema



Business Query

Show the percentage of adults in each Income Bracket, grouped by their Marriage Category

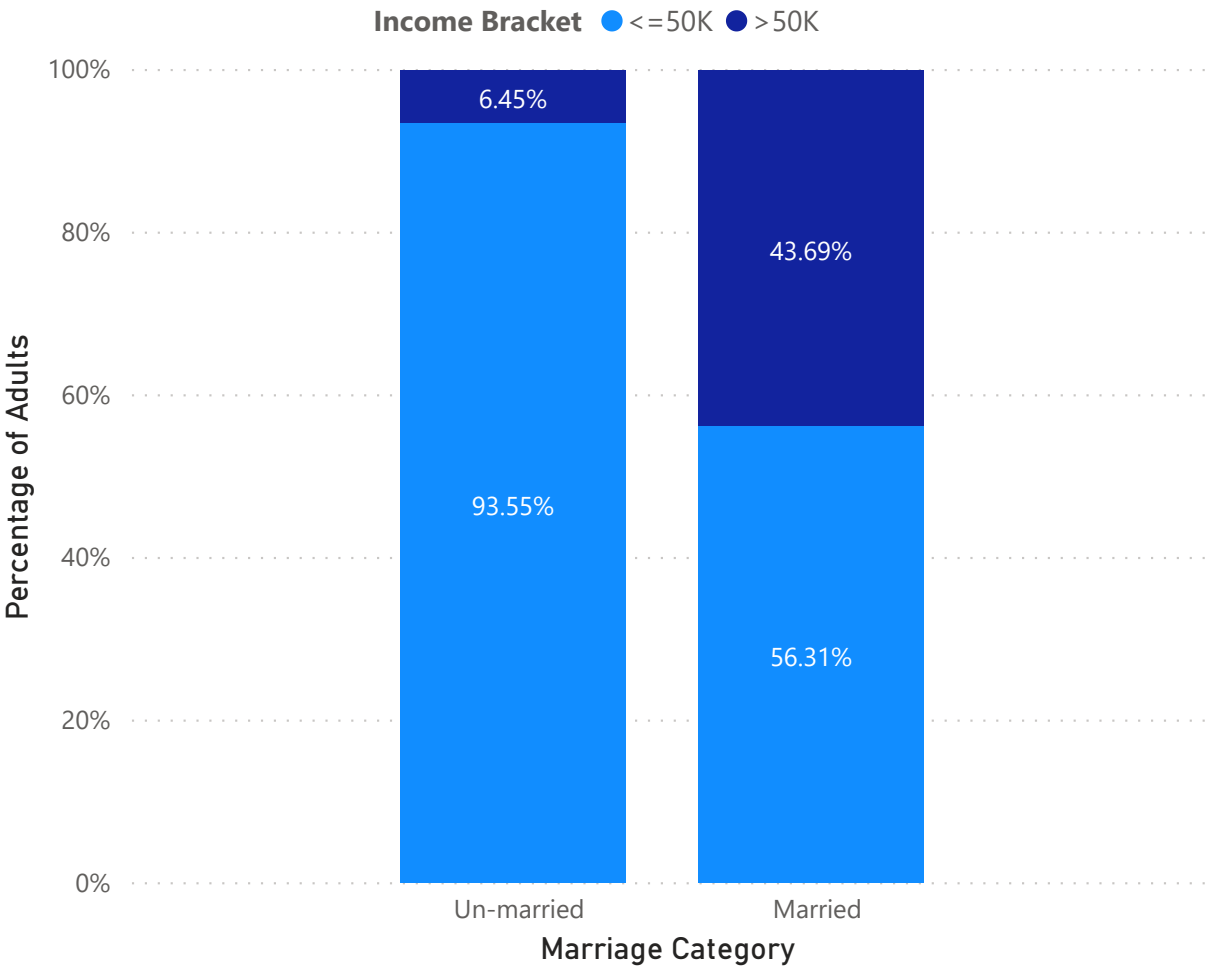
Analysis

This data shows that a much larger proportion of un-married adults earn less than or equal to \$50,000 when compared to married adults.

Of adults who are not married, 93.55% earn less than or equal to \$50,000, while only 6.45% earn more than \$50,000.

Of adults who are married, 43.69% earn more than \$50,000 while 56.31% earn less than or equal to \$50,000.

Adult Count by Marriage Category and Income Bracket



Marriage Category	<=50K	>50K	Total
Un-married	16039	1105	17144
Married	8681	6736	15417
Total	24720	7841	32561

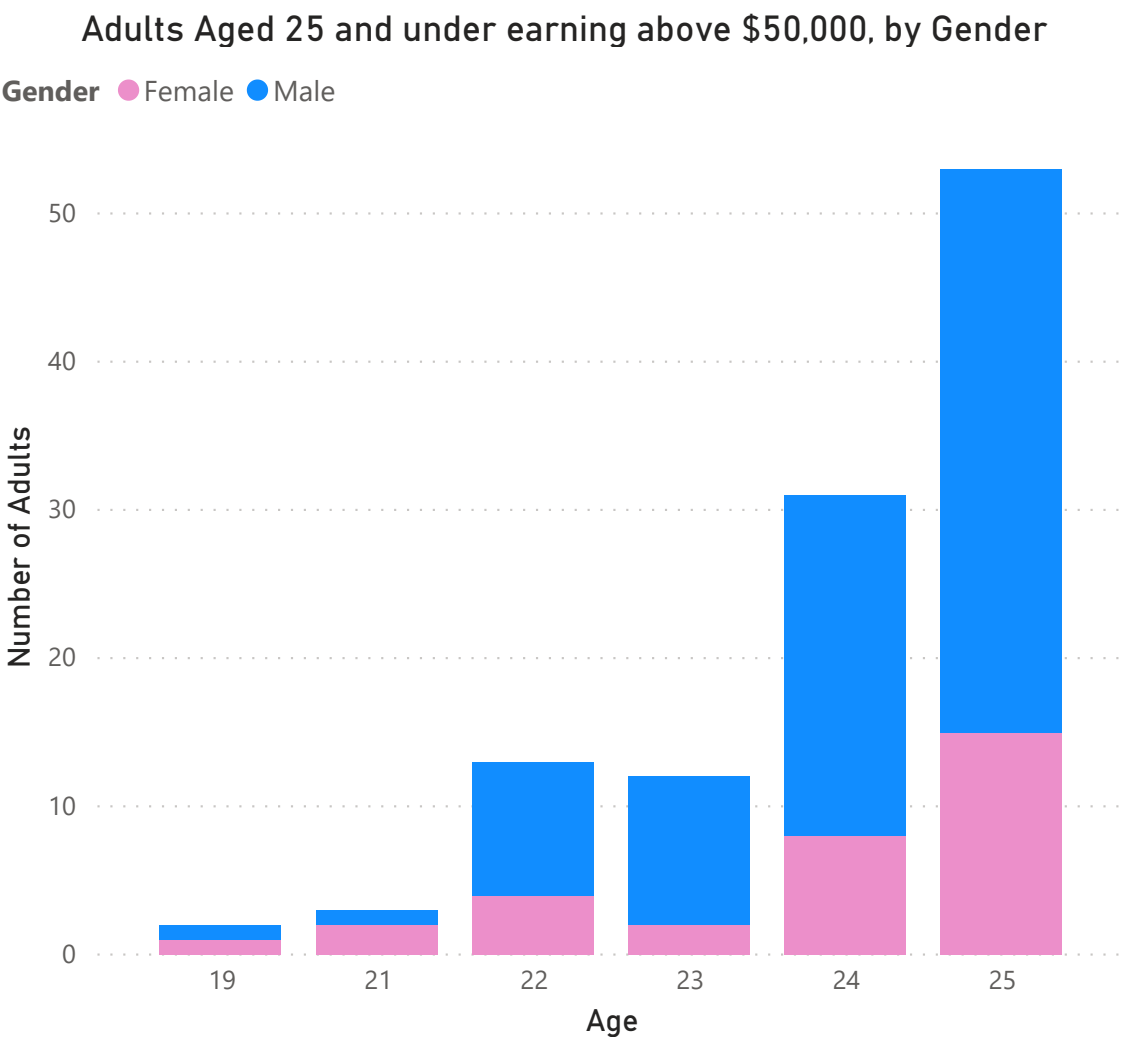
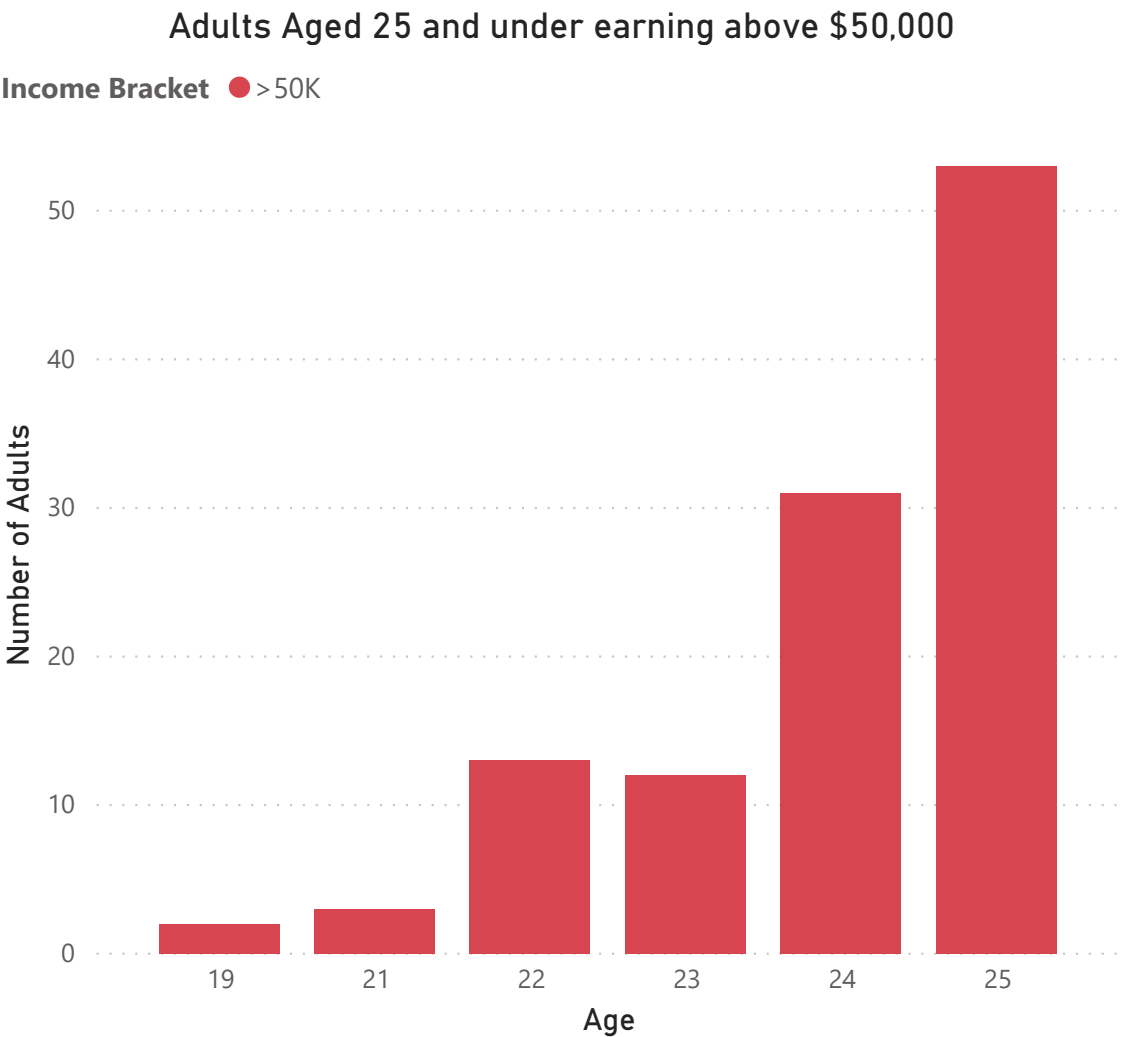
Business Query.

How many adults aged 25 and under earn more than \$50,000?
Show the number per gender

Analysis

Here we can see that people in their early 20's generally have a higher income with age, with a sharp increase at ages 24 and 25.
In this age group, there are more Males earning higher incomes than Females.

Age	Female	Male	Total
19	1	1	2
21	2	1	3
22	4	9	13
23	2	10	12
24	8	23	31
25	15	38	53
Total	32	82	114



Business Query.

Show the number of adults in each Income bracket, grouped by:

- Age Bracket
- Educational Institution they attended
- Occupation Type

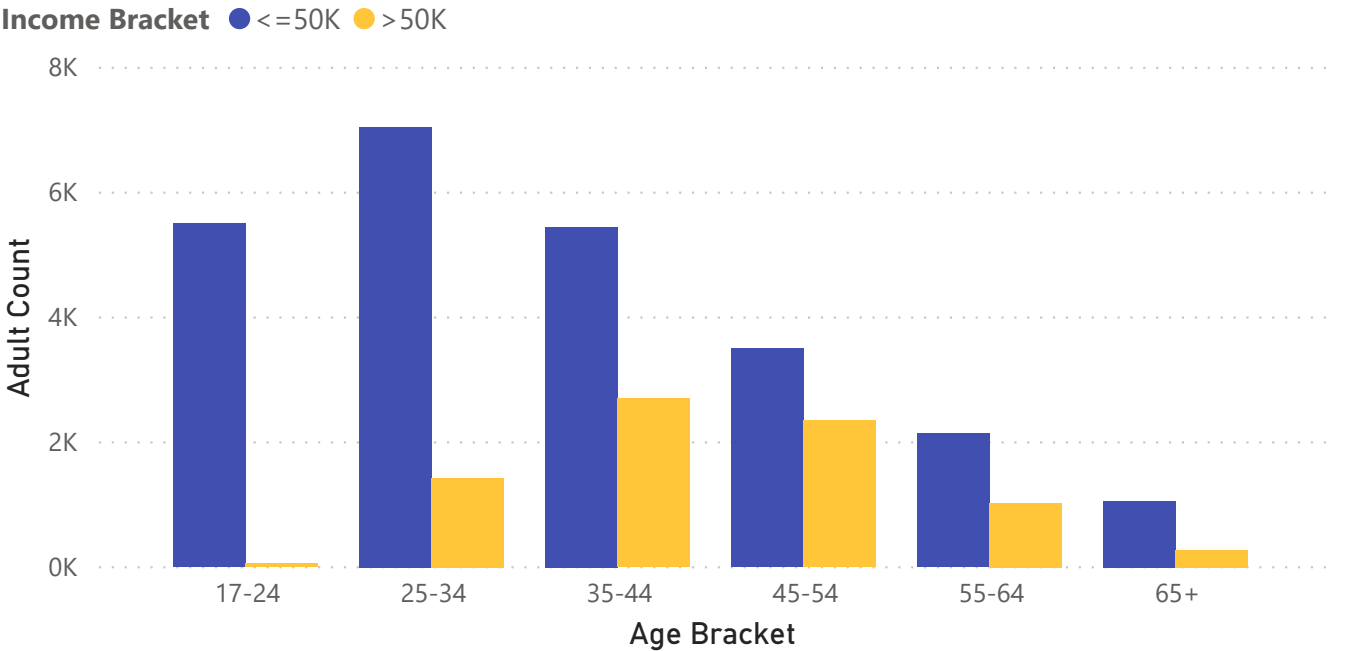
Analysis

For those who went to University, there is a rather equal distribution of people in both Income Brackets. For those with College and High school as their highest education, far less earn more than \$50,000 proportionally.

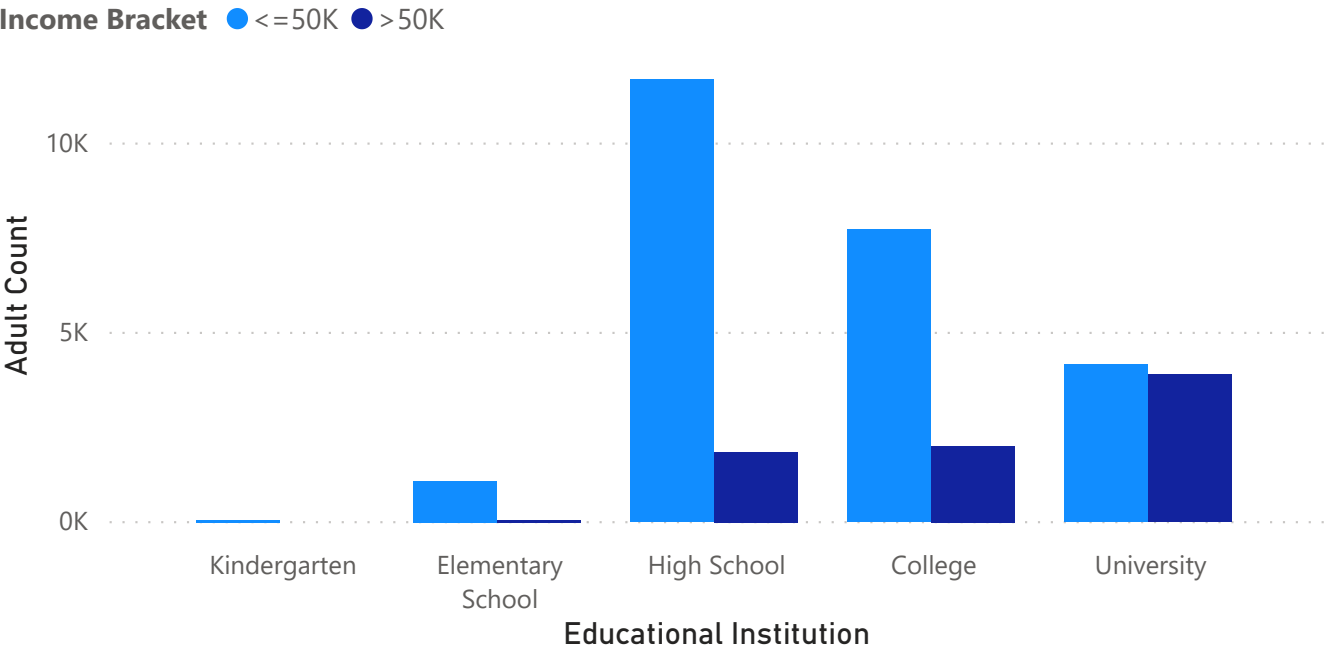
The age bracket with the most equal distribution of adults earning more than \$50K and less than \$50K is ages 45-54. In every other age bracket, the number of adults earning less than \$50K is much higher than the amount earning more than \$50K.

In terms of Occupation, the jobs with the highest number of people earning more than \$50K are Prof-specialty and Exec-managerial.

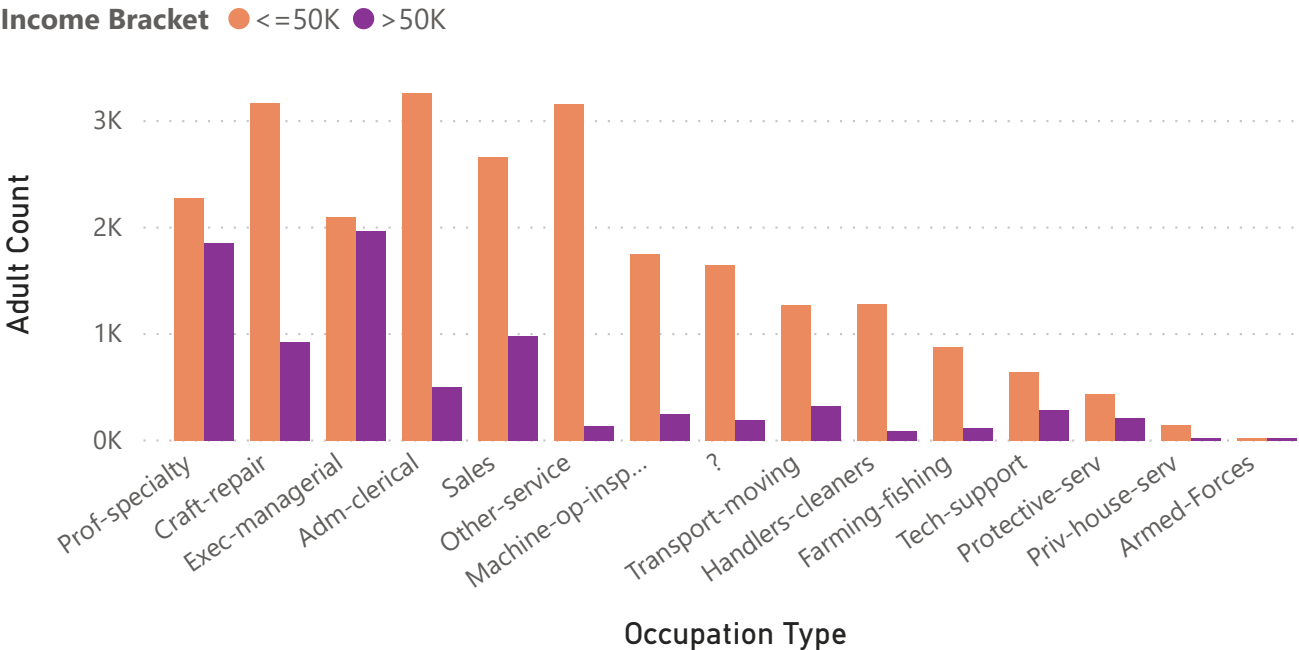
Adult Count by Age Bracket and Income Bracket



Adult Count by Educational Institution and Income Bracket



Adult Count by Occupation Type and Income Bracket

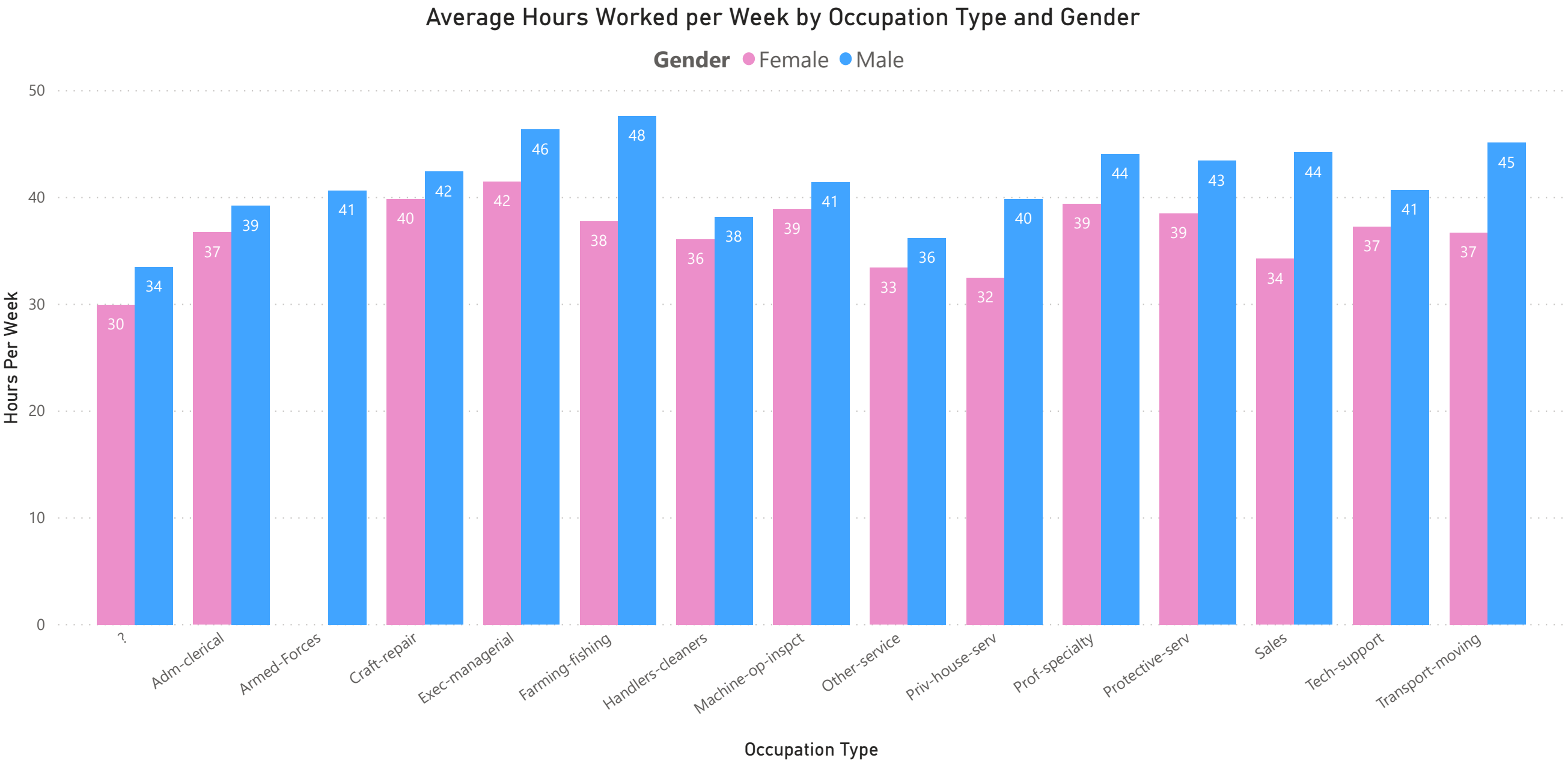


Business Query

What are the Average Number of Hours worked per Week by Occupation Type and Gender

Analysis

On average, men work more hours per week than women, and the ratio is higher in Job types such as Farming-fishing, Transport, and Sales



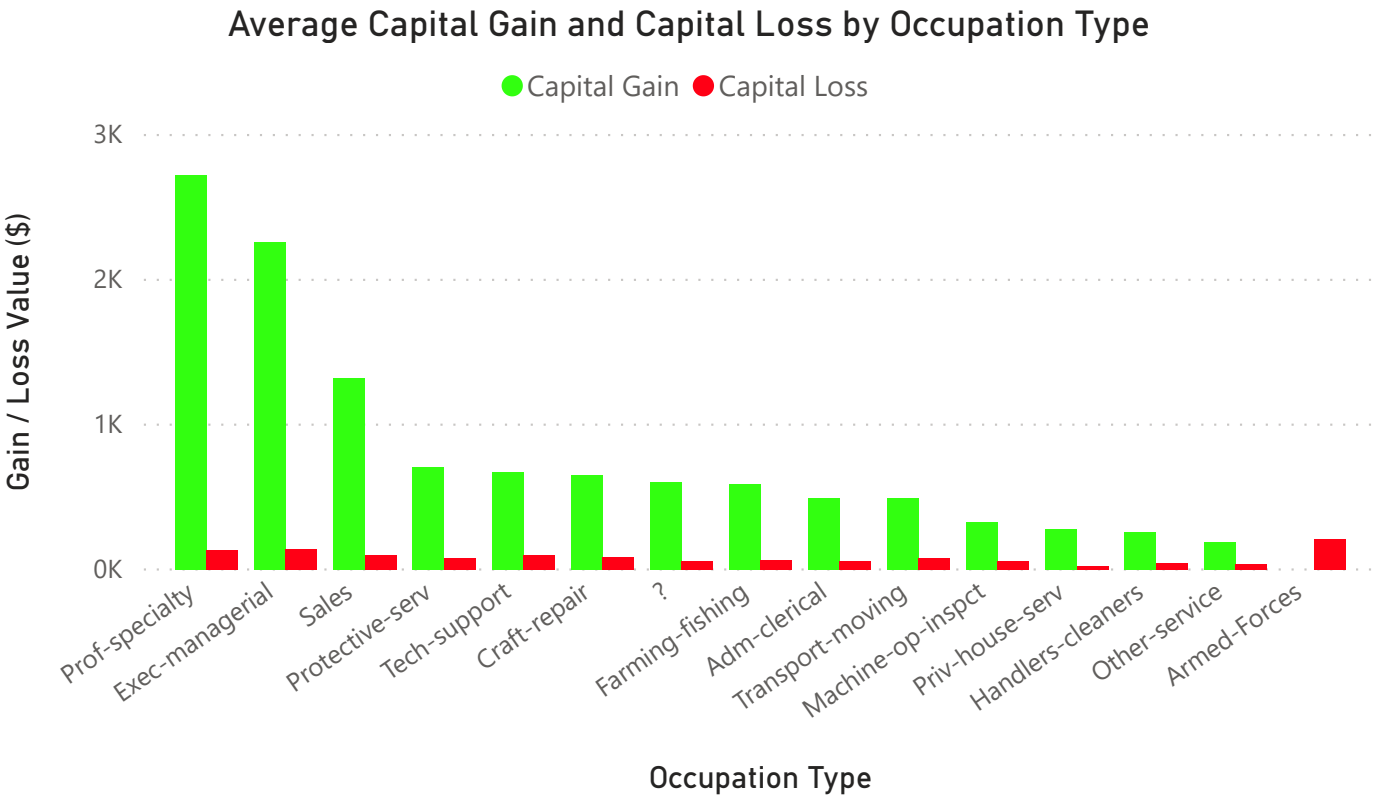
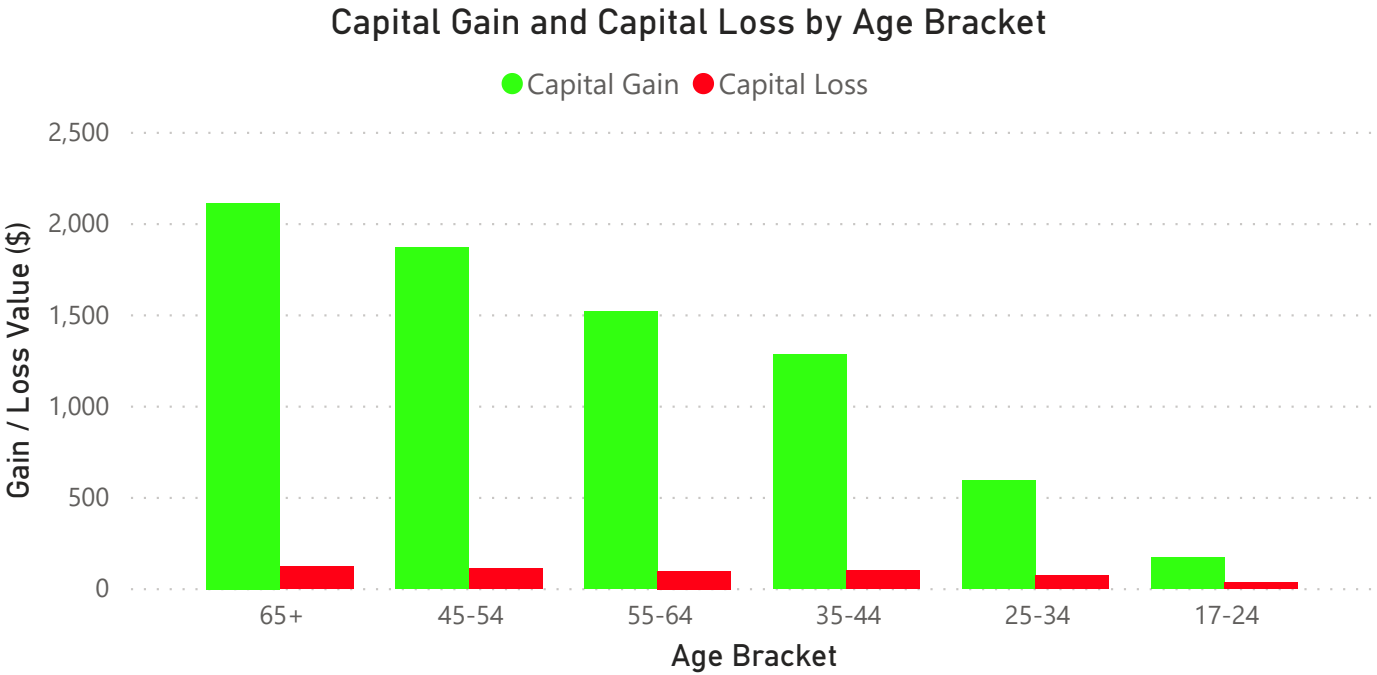
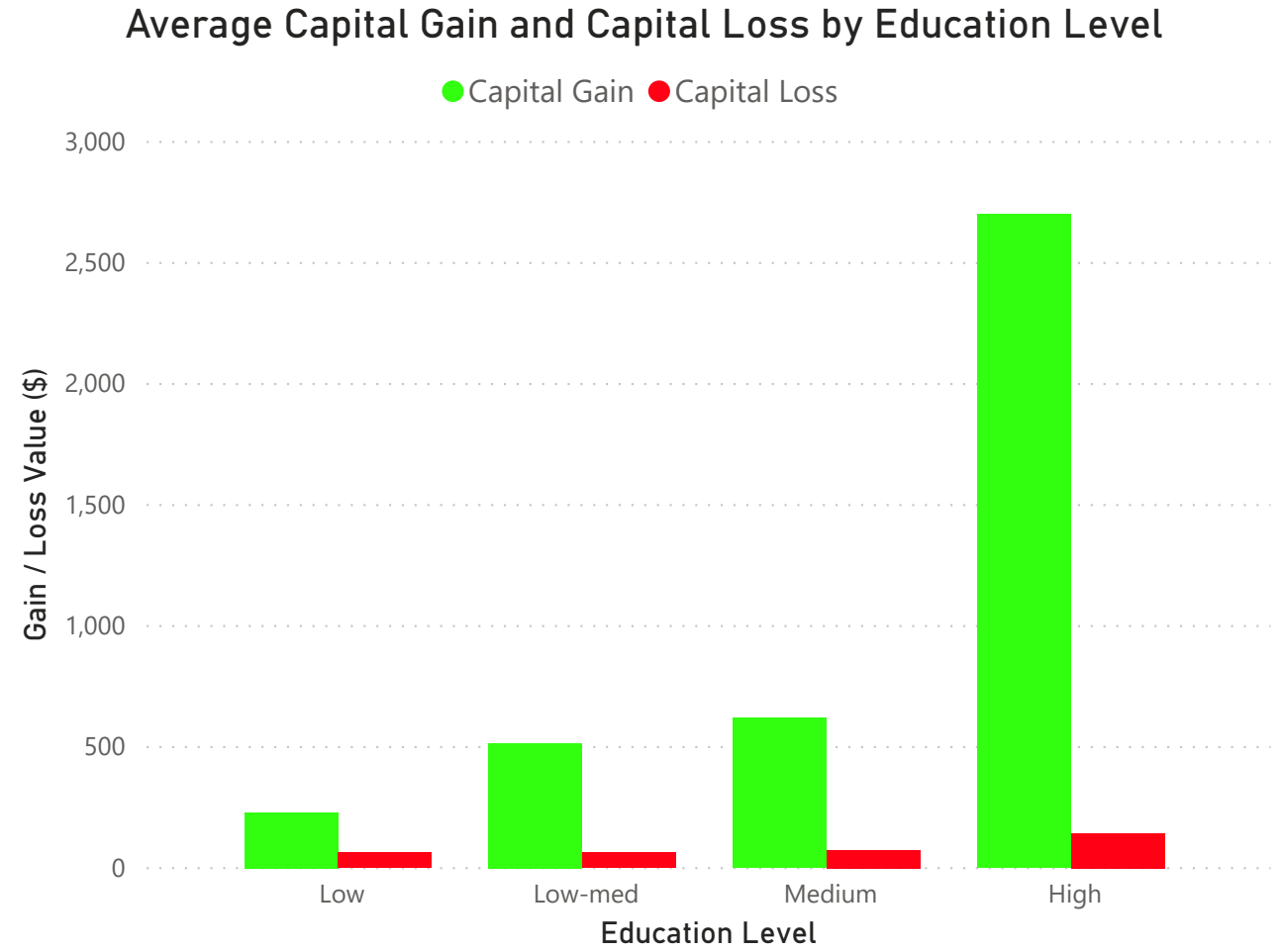
Business Query

What is the average Capital Gain and Capital Loss according to:

- Age Bracket
- Education Level
- Occupation Type

Analysis

In general, the Average Capital Gain is much higher for those with a high level of Education (University Level), with an average of \$2701.01. This is more than 4 times higher than those in the second highest Education level group, who had an average Capital Gain of \$619.87. The trend shows that Capital Gain greatly increases in higher Age groups, and for those who work in Professional-specialty, and Exec Managerial jobs, perhaps due to more knowledge and experience with investing.



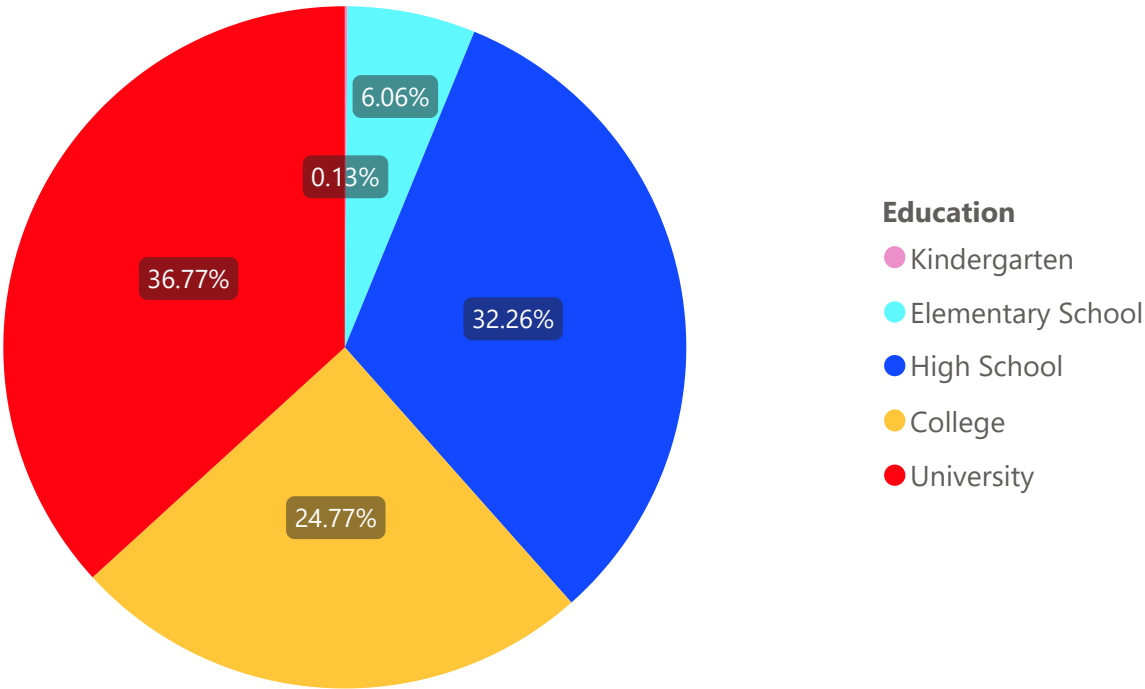
Business Query

Show the percentage of Adults per Educational Institution, categorized by whether their Native Country is classified as Developing or Developed

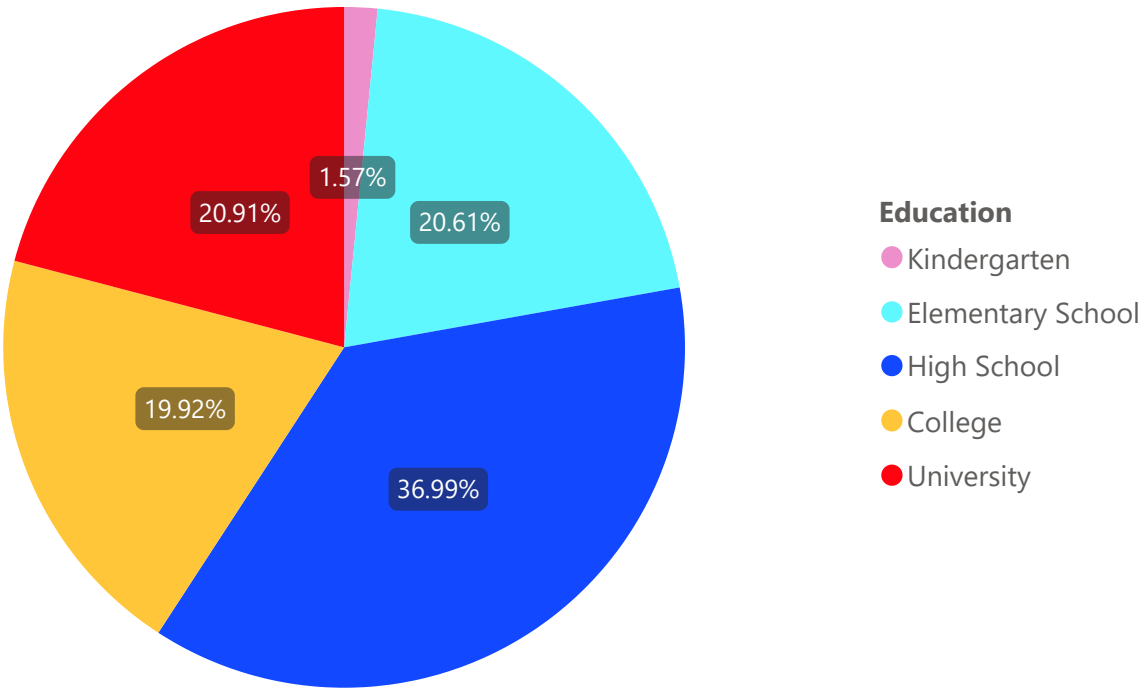
Analysis

The proportion of those whose highest Education Level is High School and College is not that different between those originating from Developed or Developing nations. The big difference is the percentage who attended University, being much higher among those whose Native Country is developed.

Education classification for Adults from Developed Nations (not including USA)



Education classification for Adults from Developing Nations



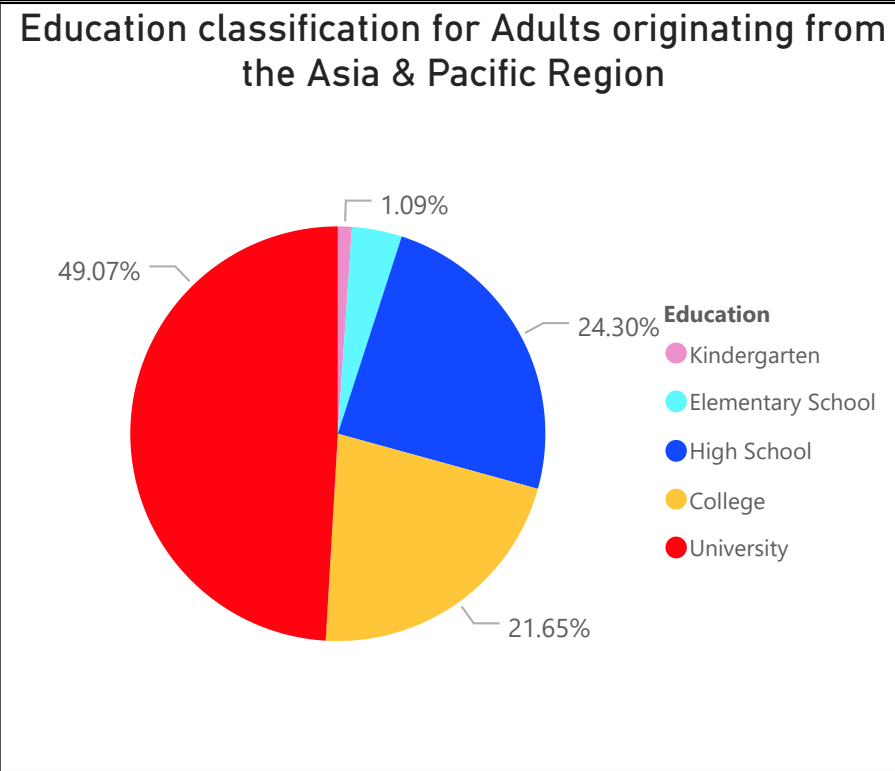
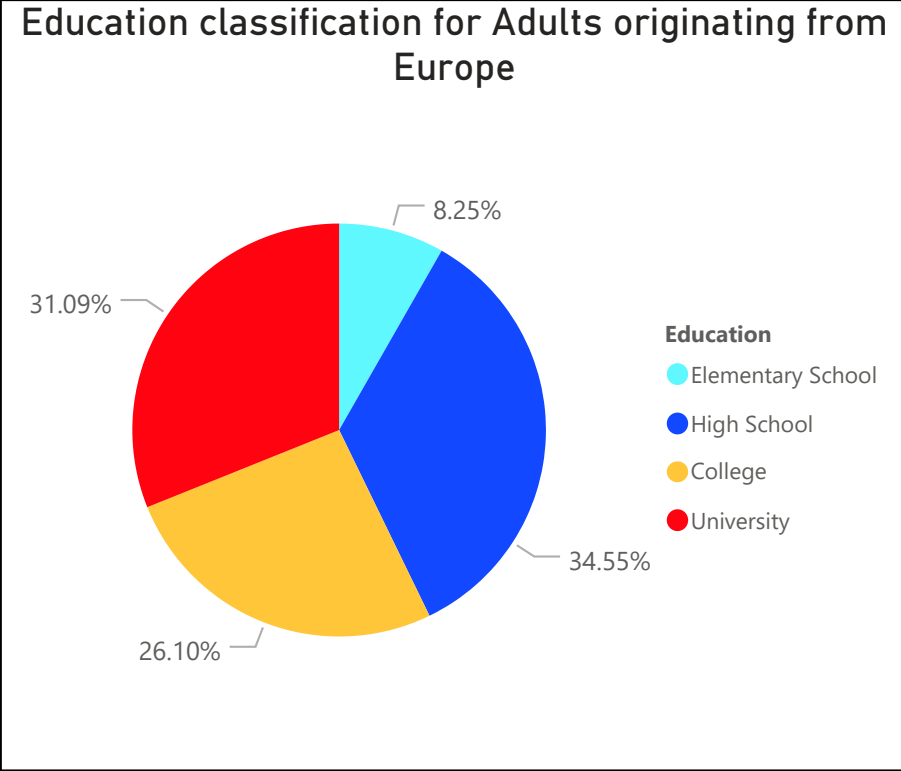
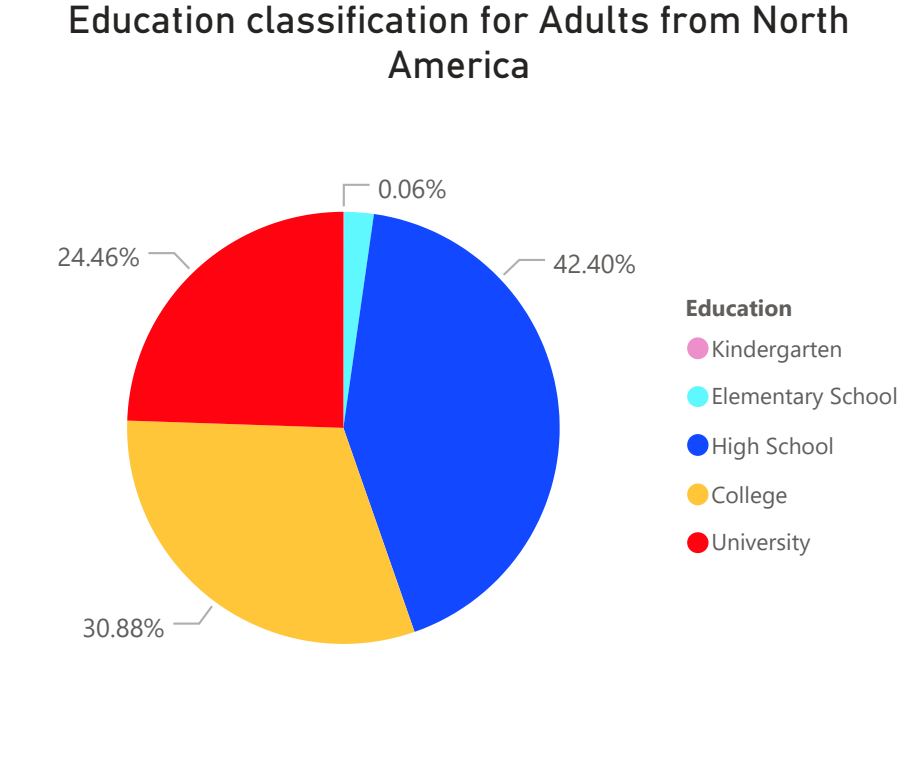
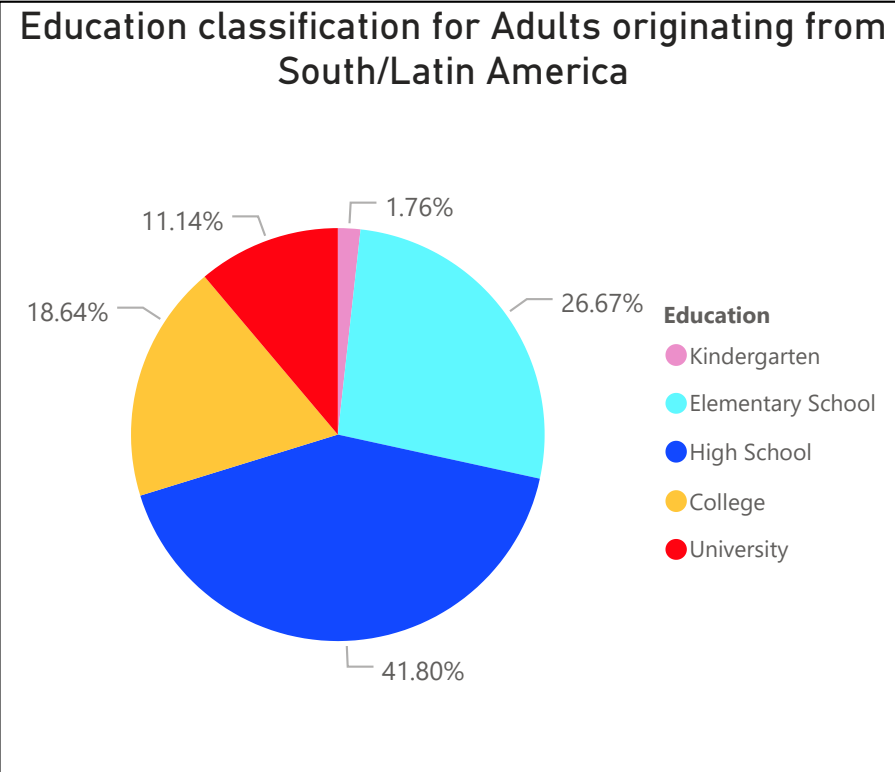
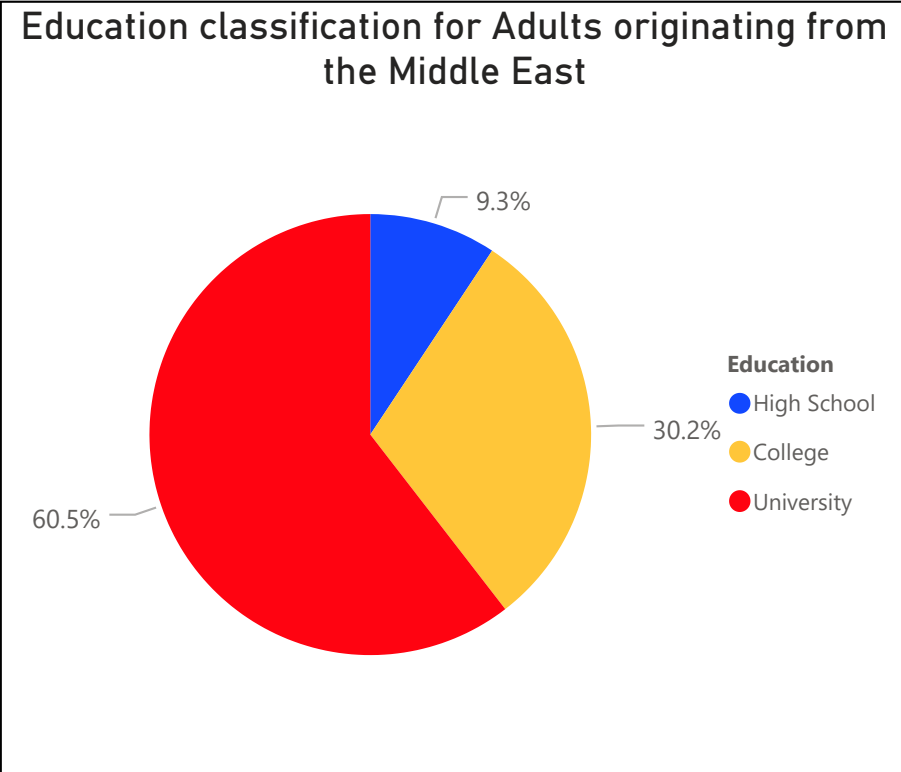
Business Query

Show the Education classification for Adults from each region

Analysis

Adults originating from the Middle East and Asia & Pacific regions have a high percentage of University attendance, compared to adults from the South/Latin America region who generally have a lower percentage of University attendance.

Adults from North America and the South/Latin America region have a similar proportion for those with High School as the highest level of education, at an approximate average of 42%.



Business Query

Show the percentage of adults per occupation type, categorized by whether they graduated from High School or not. Show which income bracket they are in

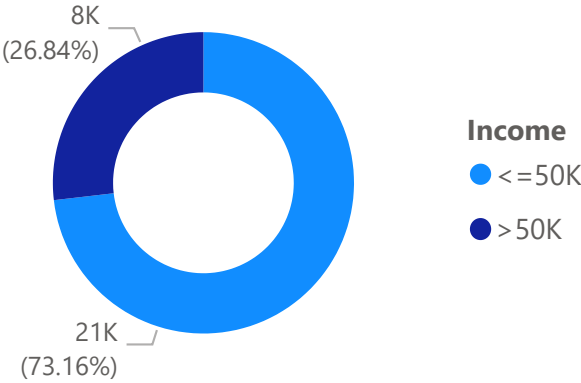
Analysis

The results confirm our expectations. There is a greater number of adults earning less than \$50,000 that did not graduate from high school, compared to those that did graduate.

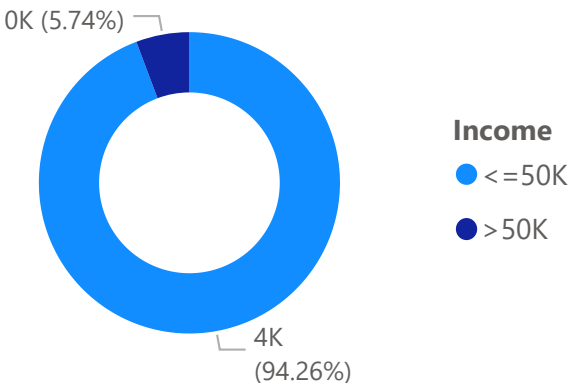
For those that did not graduate high school, the highest proportion work in jobs such as Craft-repair, Machine-op-inspecting, and Other-services.

For those who graduated high school, the highest proportion work in jobs such as Professional-specialty, Exec-managerial, and Admin-clerical.

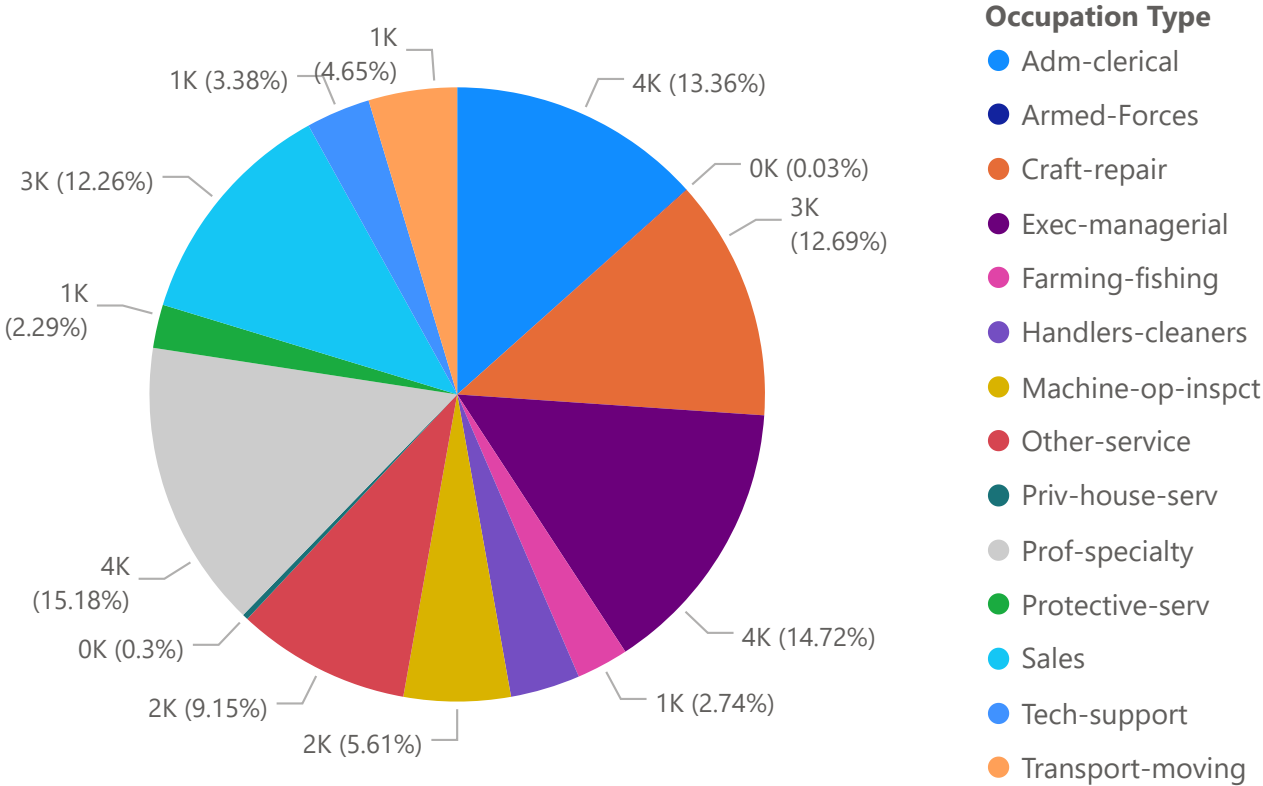
Adults per Income Bracket that Graduated High School



Adults per Income Brackets that Didn't Graduate High School



Number of Adults per Occupation Type that Graduated High School



Number of Adults per Occupation Type that Didn't Graduate High School

