Revolution Consulting

Data Modelling Employee Resignation

Harry Lloyd (s3993972)

Table of contents

Introduction:

Revolution Consulting (RC) is an IT company that offers services centred on AI machine learning and data science driven solutions. Revolution Consulting is a top performer in the field and it is an imperative for management that the company is able to attract and retain high-quality practitioners. Recently, the management team has reported increasing levels of negative feedback noting a decline in the quality of work being produced. A preliminary investigation has identified high rates of employee churn as a likely cause of the company's declining standards.

The management team at Revolution Consulting provided the Analytics team with a CSV file containing data on current and past employees of the firm. The research goal of this report was to identify factors that are likely to be contributing to employee attrition, and then identify clusters of employees who share these attributes. With this data the management team will be able to identify employees with a high-risk of resigning and develop strategies designed specifically to target these employees

This report details a preliminary exploration of the data as provided by RC's Human Resources department. This exploration provides a broad overview of the company and its employees, to better inform the preceding exploratory analysis of relationships in the data.

The following relationship analysis provides a deeper understanding of the relationships between attributes, with the intent to identify factors that impact rates of resignation and employee churn. By completing this analysis prior to modelling, the analyst team was able to omit potential sources of noise and irrelevant complexity from the data modelling process.

The third section of this report details the machine learning models (K-means & DBSCAN) used to identify potential high-risk clusters of employees. This report then evaluates a high-risk cluster to determine its characteristics and provide a recommendation to the management team to combat the employee turn-over within the grouping.

Features overview:

The data provided by Human Resources contained information on 1,470 employees.

A brief summary of the data is detailed in the table below.

Feature name	No of unique values	Type	Description
EmployeeID	1470	Nominal	A unique identification number given to every employee.
Age	43	Interval /Ratio	Age of the employee at time of survey. (18-60)
Resigned	2	Nominal	Employment status of the employee.
BusinessTravel	3	Ordinal	Categorical values (Frequent, Rarely, Non) denoting the frequency of travel undertaken by an employee.
BusinessUnit	3	Nominal	Categorical values (Sales, Consultants, Business Operations) for each business department.
EducationLevel	5	Ordinal	Categorical values (1-5) denoting employee education, with 1 being lowest value.
Gender	2	Nominal	Gender (Male/Female) of the employee
JobSatisfaction	4	Ordinal	Categorical values (1-4) indicating self-reported job satisfaction, with 1 being the lowest value.
MaritalStatus	2	Nominal	Categorical values (Single, Married, Divorced) indicating marital status at time of survey.
MonthlyIncome	1349	Interval /Ratio	Monthly income for each employee. Unclear if indicative of total remuneration or 'take home' pay.
NumCompanies Worked	10	Interval /Ratio	Number of companies an employee has worked for. Given the presence of '0' values, it is assumed this value indicates 'other' companies worked for.
OverTime	2	Ordinal	Categorical value (Yes/No) indicating if an employee has worked overtime.
PercentSalaryHike	15	Interval /Ratio	Salary increases as a percentage. Ranges from 11% to 25%.
PerformanceRating	2	Ordinal	Categorical value denoting employee performance.
AverageWeekly HoursWorked	23	Interval /Ratio	Average weekly hours worked by an employee.
TotalWorkingYears	40	Interval /Ratio	Total years an employee has been working in their lifetime
TrainingTimes LastYear	7	Interval /Ratio	Times an employee underwent a training program. Ranges from (0-6).
WorkLifeBalance	4	Ordinal	Categorical values (1-4) indicating self-reported job satisfaction, with 1 being the lowest value.
YearsAtCompany	37	Interval /Ratio	Years an employee has been employed by Revolution Consulting.
YearsInRole	19	Interval /Ratio	Years an employee has been in their current position at the firm.
YearsSinceLastPro motion	16	Interval /Ratio	Years since an employee was last promoted
YearsWithCurrMan ager	18	Interval /Ratio	Number of years an employee has worked for their current manager.

Methodology:

Overview:

The objective of this analysis was to identify key factors contributing to high rates of employee churn and attrition within Revolution Consulting. First a preliminary data exploration was carried out looking into single attributes within the data to identify relevant or causal factors. Then the team engaged in a targeted relationship analysis of identified causal factors with the intent of deepening our understanding of potential causes of employee churn. Following these evaluations, unsupervised machine learning models were employed to analyse the data. This report used the K-Means and DBSCAN machine learning models to cluster the data identified in the first two steps of this analysis.

Process:

Exploratory analysis:

In order to understand the wider context within which the data modelling would take place, the data analysis team first undertook a preliminary assessment of the data provided by Human Resources. Designed with the intent to uncover potential causal factors, this exploratory analysis looked at 20 employee attributes that were believed to have an influence on employee resignation rates. Each attribute was first assessed in isolation to identify any correlation it had with employee resignation.

Relationship analysis:

Following the initial analysis, 10 pairs of selected attributes were analysed with the intent to deepen the team's understanding of the situation, and to refine the attributes selected for the data modelling process.

Steps:

1. Import modules into the environment used for data analysis.
2. Load csv. Data into the environment and verify the data load process.
3. Preliminary DataFrame check with .info() for NaN values and dtypes.
4. Undertake preliminary data assessment (Appendix 1).
5. Undertake relationship evaluation assessment (Appendix 2).

Data Modeling:

The data modelling carried out in this report utilised unsupervised machine learning algorithms (K-means & DBSCAN) in order to cluster a selected multi-dimensional dataframe. This is a commonly used technique used to identify clusters within a dataset, allowing for a target analysis of the resulting clusters to reveal their specific attributes. In order to model the data through the K-means and the DBSCAN algorithms, the data must be organised before being fed into the system. The steps followed to organise the employee data are outlined below;

1. Drop ‘Target’ value. (‘Resigned’)
2. Drop sources of data leakage (‘EmployeeID’)
3. Drop ordinal and uncorrelated values (BusinessUnit , Gender, MaritalStatus, NumCompaniesWorked, PercentSalaryHike, PerformanceRating')

K-Means modelling:

1. Select a number of K values believed to be appropriate for modelling the data (value for K to be determined with the Within Cluster Sum of Squares method (WCSS) and the Silhouette Coefficient),
2. Test clustering for assigned K value,
3. Recalibrate K and re-model for an actionable clustering.

DBSCAN modelling:

1. Run Nearest Neighbours plot to determine epsilon (Eps) value for DBSCAN modelling,

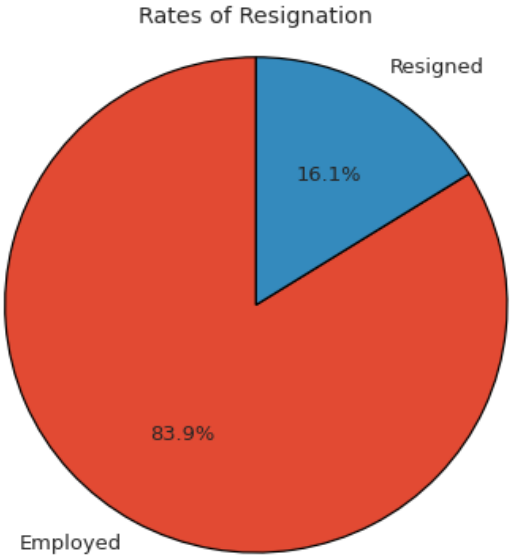
2. Test clustering for assigned Eps value,
3. Recalibrate Eps and re-model for an actionable clustering.

Evaluation strategy:

Once the machine learning models had generated an actionable set of data clusters. A single high value employee grouping was selected for an attributes analysis. This analysis contrasted the cluster's key attributes (identified in the exploratory and relationship analysis) against the general employee structure and highlighted specific metrics to target in the recommended strategy in order to minimise employee churn and resignation.

Results: Data Exploration and Modelling

Initial Data Exploration:

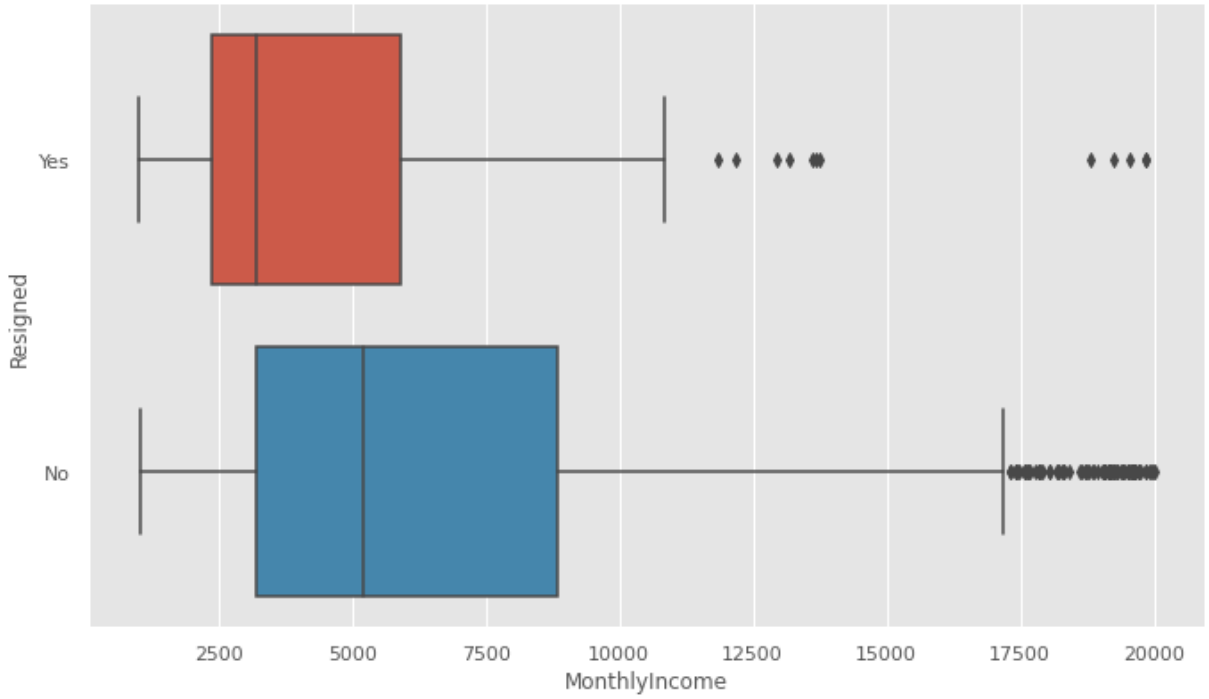


Resignation Status:

The first step of the data exploration process was an evaluation of the 'Resigned' status of the data frame. Across Revolution Consulting, 16.1% of the surveyed employees had resigned within the survey period, with 83.9% maintaining their employment (Figure 1). With this baseline set, the analysis team then worked systematically through the twenty (20) attributes against each observation to generate a more comprehensive picture of the general pattern of resignation. The following section outlines key observations made through this exploratory analysis. A complete list of these plots and observations can be found in appendix 1.

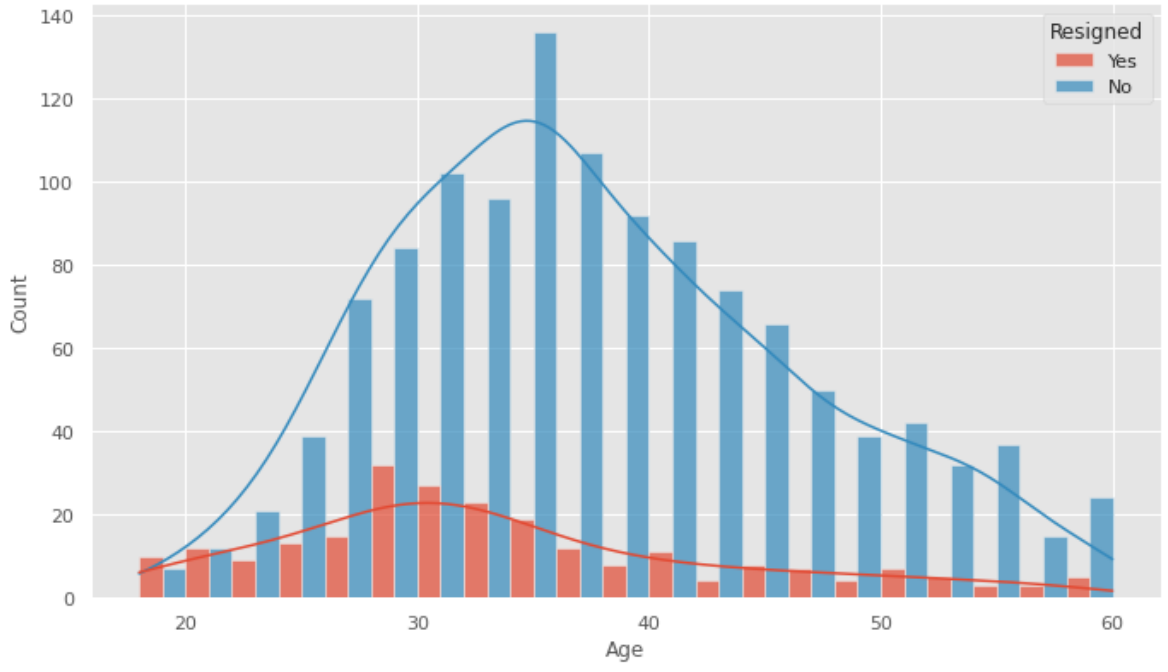
Resignation by Monthly Income:

Remuneration is widely regarded as a prime causal factor leading to employee resignation. It is perceived as a yardstick against which an employee can gauge the value of their labour to an employer and indirectly how their employer values them. This belief is reflected within the data gathered on employee monthly incomes in figure 2. Ex-employees who resigned from RC in the surveyed period received a median monthly income (approx $3,300) that was substantially lower than those who maintained their employment at RC (approx $5,100). This stark difference in income between past and present employees highlights income as a causal factor in an employee's decision to resign.



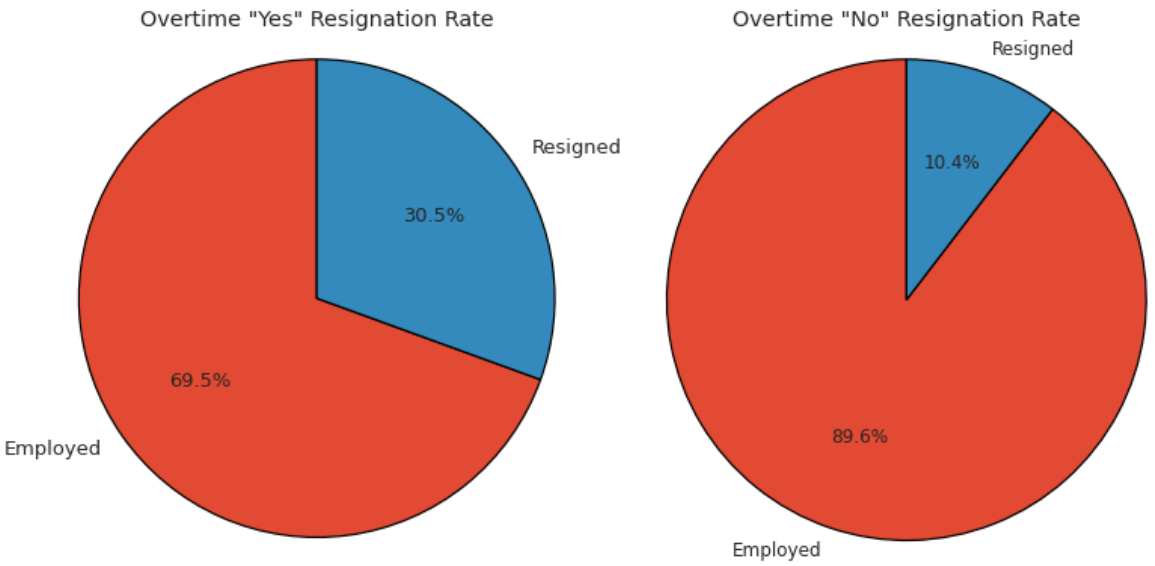
Resignation by Age:

The exploratory analysis also revealed employees under the age of 40 as a key cohort in relation to employee resignation. Peaking at around 28-30 years of age, employees who are under 40 have a substantially higher propensity to resign than their older colleagues. Interestingly, employees under 20 years of age were more likely to resign than to stay with the firm, and employees aged 20-22 had an approximately 50% chance of resigning. This churn of younger employees represents a significant long-term concern for Revolution Consulting. Younger employees are more likely to have received contemporary training and education prior to commencing their tenure and have the potential to modernise practices at the firm. These employees are also likely to require a higher degree of training and exposure, losing these employees after spending resources on their development is a poor outcome for RC.



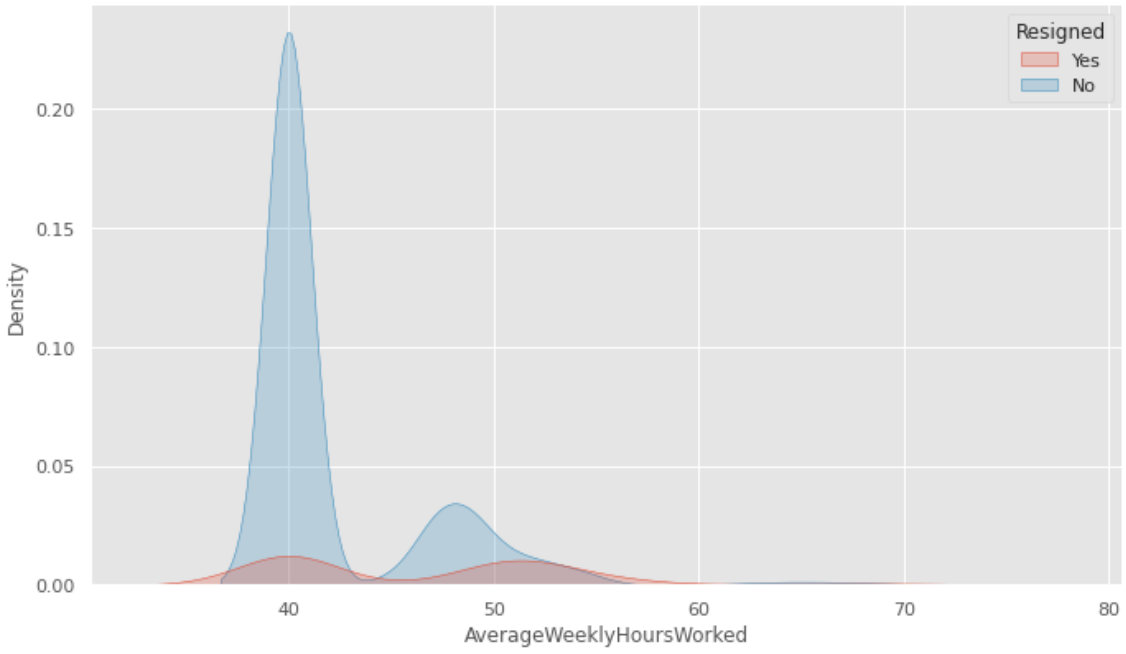
Resignation by OverTIme:

Another cohort of concern identified in the exploratory analysis was employees who recorded having worked overtime within the survey period with 30.5% of these employees resigning. When contrasted against those who were not required to work overtime, with a 10.4% rate of resignation, this makes a substantial increase in the likelihood of an employee to resign. When expressed as an approximate fraction, this is an increase from 1 in 10 employees resigning, to 1 in 3.

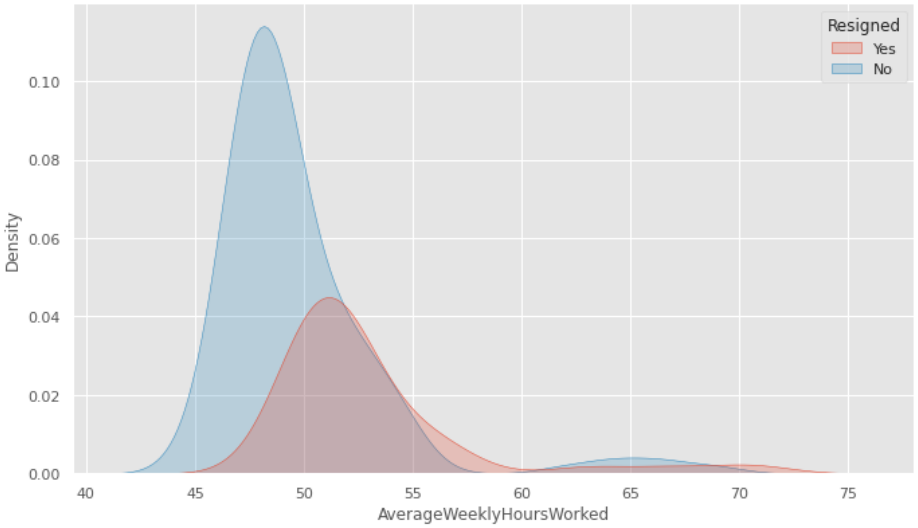


Resignation by Average Weekly Hours Worked:

The average weekly hours worked by an employee was also found through the initial data exploration to have a notable impact upon rates of resignation for Revolution Consulting's employees.For the majority of RC's employees, the average working week is 40 hours (Figure 5), there is however a considerable cohort of individuals who routinely work above this threshold.



Interestingly, if an individual is required to work up to 45 hours a week, the additional workload has little influence on their likelihood to resign. However, when employees report working from approximately 50 hours to 60 hours a week, the likelihood they resign is substantially elevated (Figure 6).



Data Relationship Exploration:

Following the initial data exploration four (4) key attributes were identified as having a marked influence upon employee resignation; Monthly Income, Age, OverTime, Average Weekly Hours Worked. These four key attributes were then used as a launch pad into an in-depth exploration of relationships between the attributes provided in the data set.

Monthly Income / Average Weekly Hours Worked (AWHW):

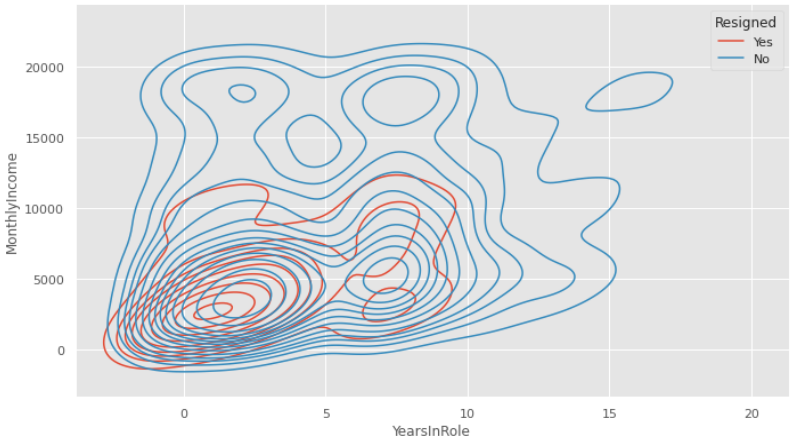
Based on the previous findings for Monthly Income and AWHW, the analysis team hypothesised that as an individual worked longer hours, they would expect to be financially rewarded as a result, and those who were not, would likely be the employees who resign.

Figure 7 illustrates Monthly Income as a function of AWHW for employees and reveals two key "at risk" clusters of employees. The first of these clusters is those employees working 40 hour weeks and receiving a Monthly Income of less than approximately $11,250. While it is difficult to discern from figure 7 alone the likelihood that employees falling into this grouping will resign, the plot

demonstrates a clear relationship between the two attributes and their influence on an employee's decision to resign. The second cluster identified in figure 7 comprises those employees working approximately 49 to 57 hours a week and earning less than $11,250. Akin to the first cluster identified, this relationship between hours worked and remuneration has a clear influence on an employees propensity to resign. It is important to note in this second cluster, that as Monthly Income decreases the density of observations increases, with a clear concentration around employees earning $2,500 a month.

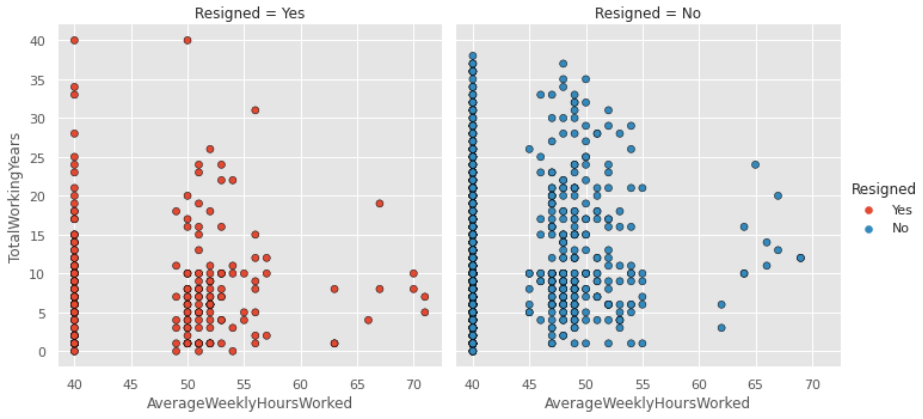
Monthly Income / Years In Role:

The analysis team hypothesised that employees who believe they are underpaid for their experience in a given role would be another cohort likely to resign. However the data showed that once an employee had over 10 years in a role the likelihood of their resignation dramatically decreased. In fact, the employees who were most likely to resign were those with under 5 years in a role who were earning less than $7,500 a month. This pattern of resignation intensifies as YearsInRole and MonthlyIncome decrease.



Total Working Years (TWY) / Average Weekly Hours Worked (AWHW):

In a continued investigation into the influence of AWHW, the analysis team paired the attribute with the reported Total Working Years of employees. The analysis team hypothesised that young employees at the start of their careers would feel overworked if they are extended beyond an average of 40 hours a week too early in their careers and would likely resign as a result. Figure 9 demonstrates that the team's hypothesis was for the most part accurate, with a distinct cluster of resigned employees who reported less than 13 Total Working Years that were concurrently working between 49 to 57 hours a week. Equally important, once employees had over 15 Total Working Years reported, the proportion who resigned substantially decreased.



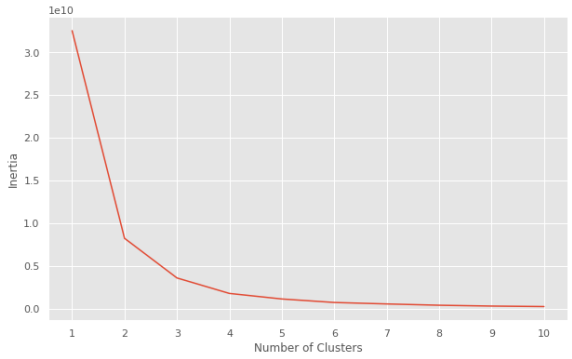
Data modelling:

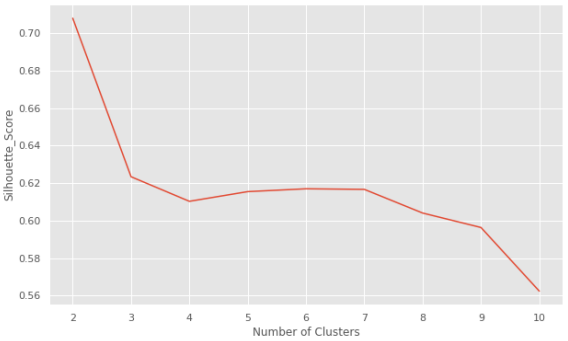
The data used within the K-Means and the DBSCAN modelling techniques only included attributes found within the exploratory and relationship analysis to have an influence on rates of resignation. For example, 16.1% of employees with a performance rating of 3 resigned, compared to 16.4% of employees with a rating of 4. From this observation it is clear that performance rating has little to no effect on the choice to resign, and as such it is dropped from the modelling data. Nominal data has also been removed.

Model 1: K-Means Cluster

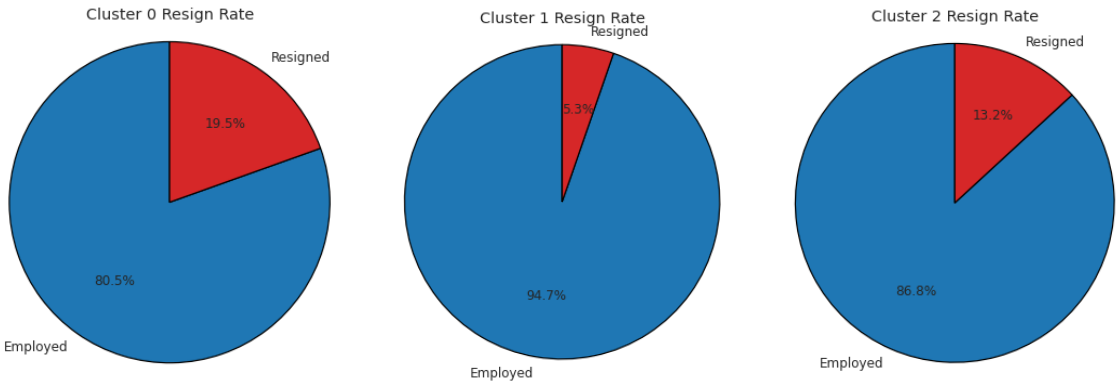
The analysis team employed the K-Means clustering technique to evaluate two selected data frames for meaningful employee clusters. The first dataframe included 14 attributes deemed to be significant and are as follows; Age, BusinessTravel, EducationLevel, JobSatisfaction, MonthlyIncome, OverTIme, AWHW, TotalWorkingYears, TrainingTimesLastYear, WorkLifeBalance, YearsAtCompany, YearsInRole, YearsSinceLastPromotion, YearsWithCurrManager.

In order to determine an appropriate K value to run through the model, a Within Cluster Sum of Square (WCSS) and a Silhouette Coefficient were graphed and have been provided below. When appraising a WCSS score, the “elbow” within the chart is used to denote an appropriate value for ‘K’, in this instance 3. This value is confirmed with a silhouette score of 0.62





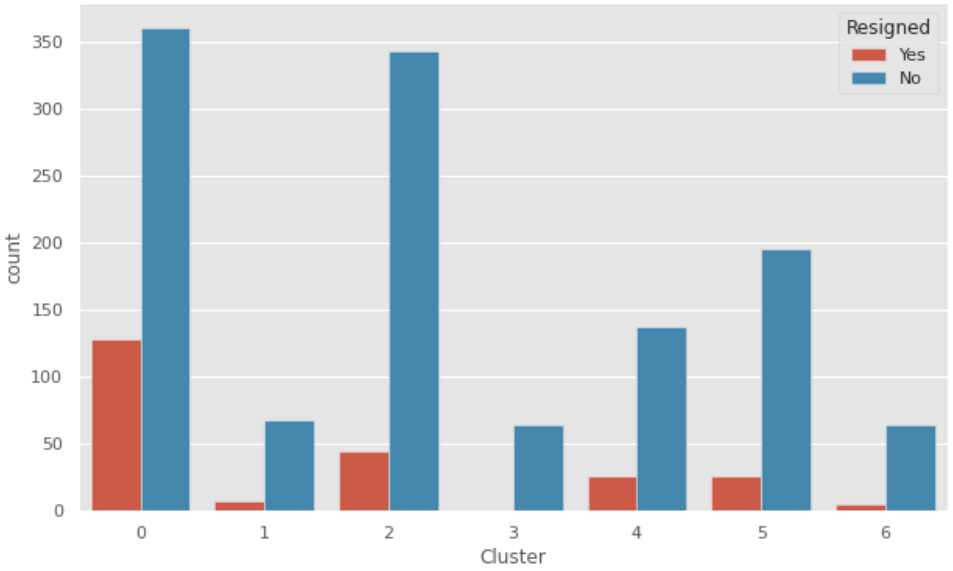
When the K-means machine learning algorithm was run however the clustering result did not provide meaningful or actionable data. The clustering result provided insignificant deviation from the general rate of resignation of 16.1% illustrated in figure 1, with Cluster 0 only increasing by 3.4%



Despite 3 appearing to be the most appropriate k value for model clustering, the intent of this report is to provide recommendations to the management team in order to reduce employee turnover. The above modelling result as such does not suit the requirements of management and the model was reconfigured and re-run.

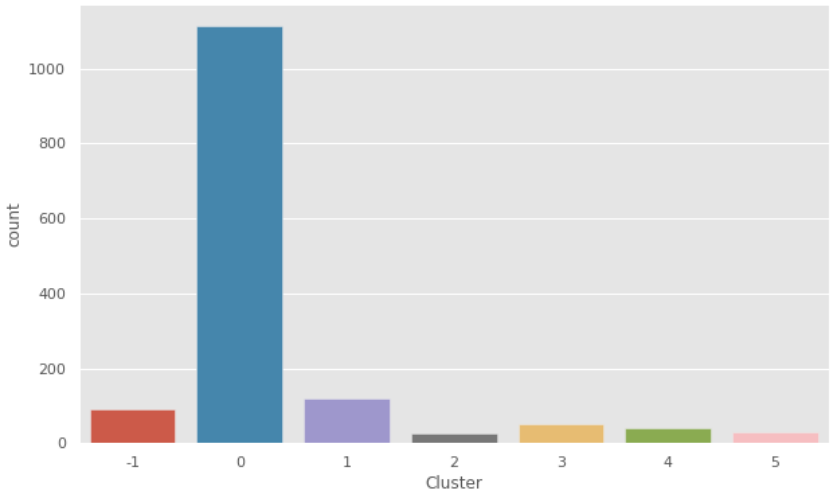
Model 2: K-Means Cluster

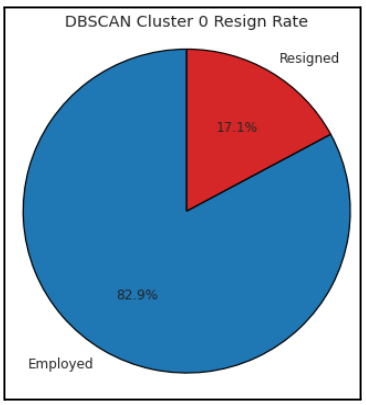
The second appropriate k value identified in the Silhouette Score plot was 7, with a silhouette score of slightly lower than 0.62 as run in the previous model when k = 3. This second clustering model produced a series with a greater degree of significance, however only cluster 0 was deemed by the analysis team to be of merit due to its size and 'Resigned' rate of 26.2% (Cluster ratios in Appendix 3). Cluster 3 is also of note with a 0% Resignation rate.



Model 3: DBSCAN (Eps 200)

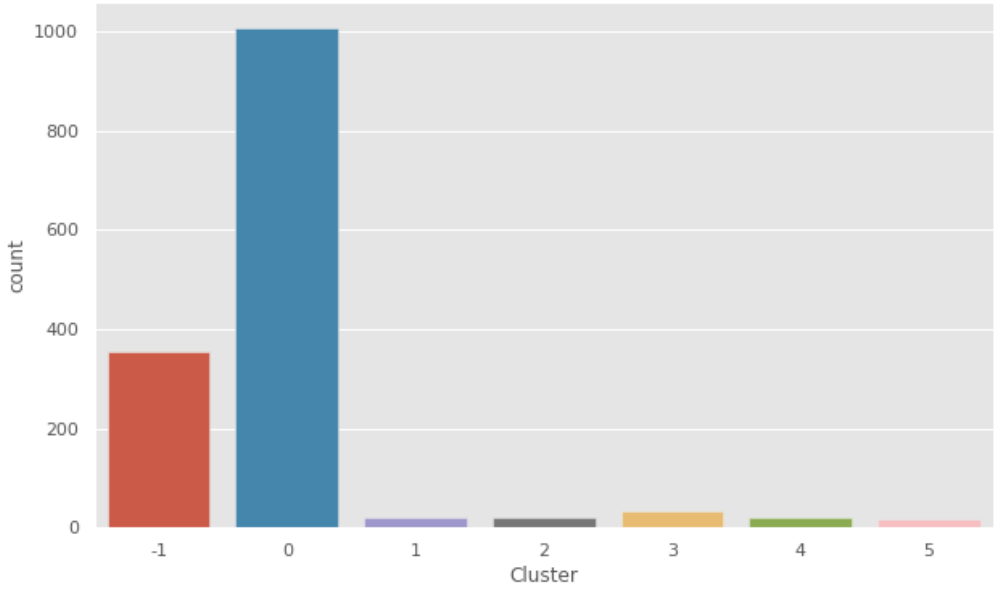
The DBSCAN method was implemented as a means to strengthen the results of the data modelling process. DBSCAN modelling clusters data points based on their relative density to one another and unlike k-means modelling, self-identifies the number of clusters through a calculation of density as a function of distance (epsilon) and a minimum number of neighbours (MinPts). The variable assigned to MinPts is usually set to the number of attributes, in this case 14, plus 1, giving a final value of MinPts = 15. Epsilon is calculated in a similar fashion to WCSS, and is plotted on the following page. In a similar fashion to selecting for K with the WCSS value, epsilon is decided by locating the 'elbow' in the plot. In this instance an epsilon of 200 was selected. When the model was run however the clustering result was ineffectual, with the vast majority of observations evaluated into a single cluster (cluster 0) with a resignation rate (17.1%) similar to the firm wide value (16.1%) in figure 1.





Model 4: DBSCAN

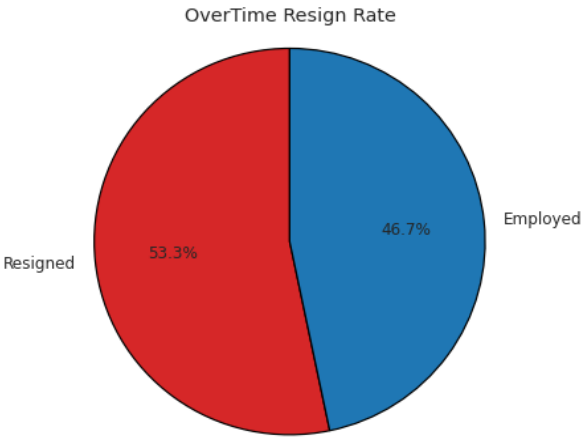
Following the over-clustering of Model 3, the DBSCAN algorithm was recalibrated and run with a decreased epsilon value of 100. This change attempted to raise the observation density required before the DBSCAN model attributed a cluster value to an observation. This change was made in an attempt to coerce the DBSCAN algorithm into generating a productive clustering outcome model to aid in our investigation. In spite of this change, the resulting model was dominated by a single cluster and the algorithm classified a large number of variables as 'noise'.



Discussion:

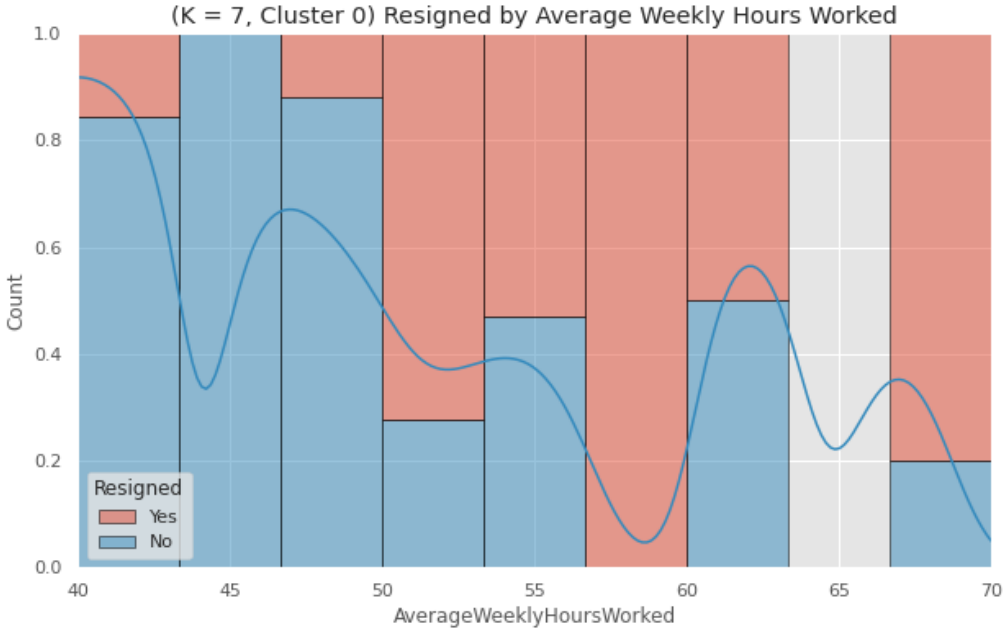
Of the two approaches implemented in this analysis, the K-means model was able to produce a meaningful set of clusters, with the DBSCAN model proving ineffective. A likely cause for this error was the high dimensionality of the dataframe utilised in the models, and given an extended project timeline the dimensionality of the data could be refined to produce a clearer model of clusters. Out of the 4 models run in this analysis, Model 2 (K-means) produced the most applicable set of clusters upon which the following recommendations are based. Cluster 0 is a key cohort for the management team to address and its general characteristics are laid out below. It captures approximately 33% of the total employee structure and has a rate of resignation of 26.2%, or just over 1 in 4 employees.

Overtime: has been identified throughout this report as a major factor contributing to the likelihood of employee resignation. Within cluster 0, If an employee worked overtime they were more likely to resign, than to continue their tenure with the company. Reducing the % of employees who are required to work overtime should be a core priority for the management team.



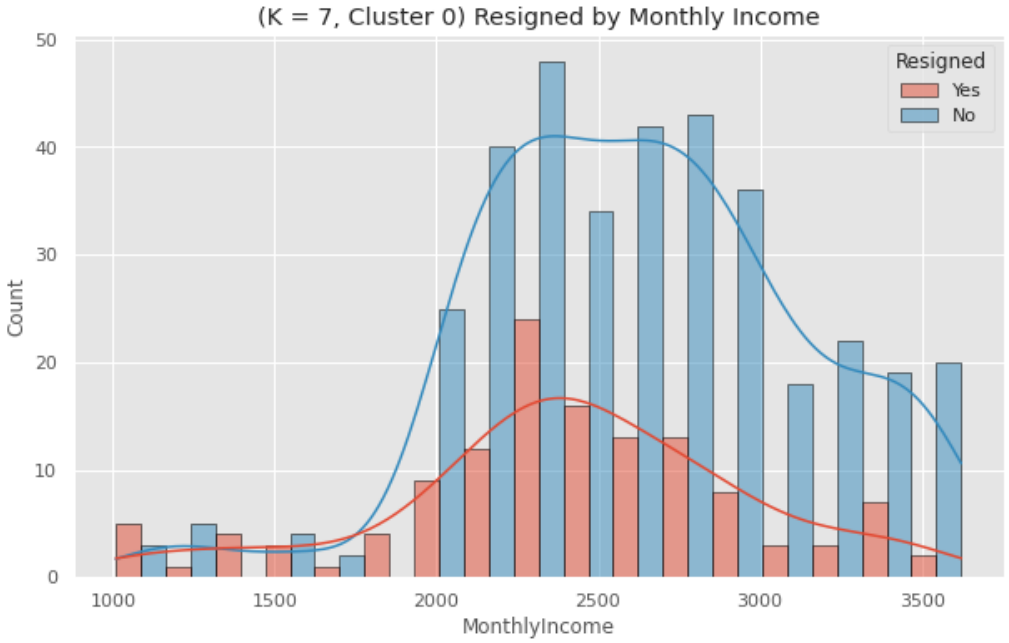
Average Weekly Hours Worked (AWHW):

AWHW is inherently tied to Overtime, however it provides more detailed insights than the binary attribute it falls under. Employees who complete over 40 AWHW but remain under 45, have a far higher retention rate than those who exceed 45 hours.



Monthly Income:

Monthly Income is another core attribute contributing to employee resignation. Within cluster 0, employee's monthly income ranged from $1,009 to $,3622. Employees who earn under approximately $1750 a month are more likely to resign than to continue their tenure with the company. Reducing the number of employees taking home less than $1,750 a month will reduce the rate of employee churn amongst low level employees.



Recommendation:

The data analysis team proposes the introduction of an employee incentive program tailored to target the three attributes summaries above; OverTime, Average Weekly Hours Worked, & Monthly Income. The program would allow low income employees earning under $3,800 a month to apply to work an extra 5 hours across the working week in return for an elevated hourly rate across those extra hours. By encouraging employees to share the load of overtime hours, the program will directly reduce the total number of employees working over the critical 45 average hours per week. As mentioned previously, this increase of 5 hours, up to an average of 45 hours a week, has no impact on rates of resignation. By directing this incentive at employees earning below $3,800, the program concurrently targets those who are earning below the at-risk monthly income threshold of $3,622 and offers them remuneration above their standard rate. This is designed to lift their monthly income as high as possible without crossing the high risk 45 hour mark. Additionally, Rather than seeing an increase in total overtime worked at the firm, this program is designed to redistribute the required labour and as such will not place additional financial stressors on the firm.

Conclusion:

This report steps through three core stages of the data science pipeline. First evaluating each attribute in an exploratory analysis, identifying target variables within the survey data (MonthlyIncome, Age, OverTime, AverageWeeklyHoursWorked) Following this stage, these identified attributes were used as a launching pad to undertake an analysis into the complex relationships between employee attributes and how these influence employee churn and attrition. Finally the report produced 4 unsupervised machine learning models (2 K-Means & 2 DBSCAN), to cluster the data into meaningful groups, allowing for the generation and recommendation of an incentives program designed to retain employees.

***Limitations**:* This report utilised the K-Means and the DBSCAN unsupervised machine learning models to interpret the complex data provided by Human Resources. These models, while effective, can be vastly improved through the implementation of various scaling techniques. Additionally, other machine learning techniques may have been better suited to the provided data, however an investigation into these alternate machine learning models was outside of the scope of this report.

Additional Comments: Throughout the plotting process the colour schemes of Seaborn and Matplotlib repeatedly inverted, resulting in colouring against usual convention. ie. Blue for negative values ie ‘Resigned’ and Red for positive. This occurrence was also inconsistent, with some plots generating appropriately whilst others did not. The colour schemes used throughout this document should be amended to ensure they are consistent and in keeping with convention prior to the circulation of this document.

References:

ChartExpo (2023) How to Read a Box Plot Chart? Easy-to-follow Steps, ChartExpo, accessed 7 February 2023. https://chartexpo.com/blog/how-to-read-a-box-plot

Corey Schafer (2020) Matplotlib Tutorial (Part 3): Pie Charts, Youtube, accessed 5 February 2023. https://www.youtube.com/watch?v=MPiz50TsyF0

Datagy (2022) Seaborn histplot – Creating Histograms in Seaborn, Datagy.io, accessed 7 February 2023. https://datagy.io/seaborn-histplot/

MatPlotlib, n.d., MatPlotlib 3.6.3 Documentation, MatPlotlib, accessed 7 February 2023. https://matplotlib.org/stable/index.html

RMIT University, n.d, Easy Cite referencing guide, RMIT Library, accessed 11 February 2023. https://www.lib.rmit.edu.au/easy-cite/?styleguide=styleguide-1#stn-5#subtype-36

Seaborn (2022) Seaborn.catplot, Seaborn, accessed 7 February 2023. https://seaborn.pydata.org/generated/seaborn.lineplot.html

Seaborn (2022) Seaborn.histplot, Seaborn, accessed 7 February 2023. https://seaborn.pydata.org/generated/seaborn.histplot.html

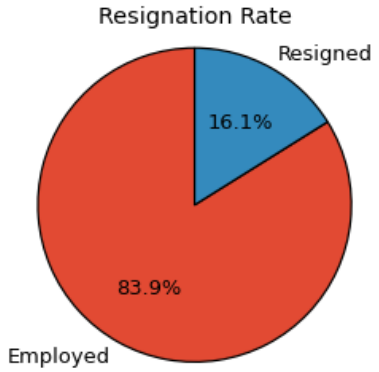
Seaborn (2022) Seaborn.lineplot, Seaborn, accessed 7 February 2023. *https://seaborn.pydata.org/generated/seaborn.lineplot.html*

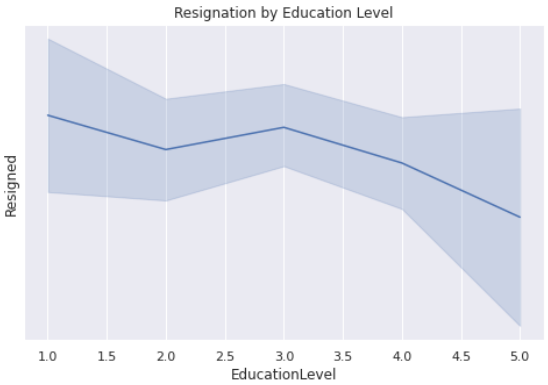
Seaborn (2022) Seaborn.boxplot, Seaborn, accessed 7 February 2023. https://seaborn.pydata.org/generated/seaborn.boxplot.html

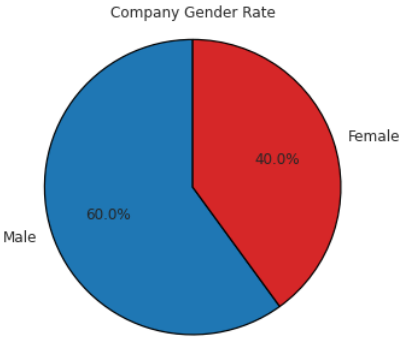
Seaborn (2022) Seaborn.boxplot, Seaborn, accessed 7 February 2023. *https://seaborn.pydata.org/generated/seaborn.stripplot.html#seaborn.stripplot*

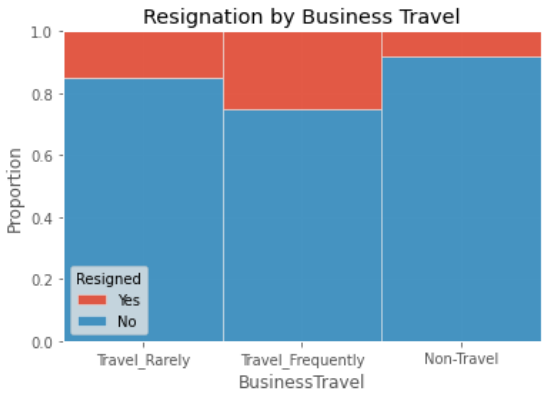
Statology (2023) 'How to Change Axis Labels on a Seaborn Plot (With Examples)', Statology, accessed 10 February 2023. https://www.statology.org/seaborn-axis-labels/

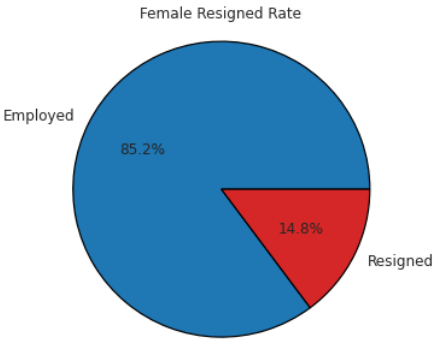
Appendix 1: Exploratory Analysis

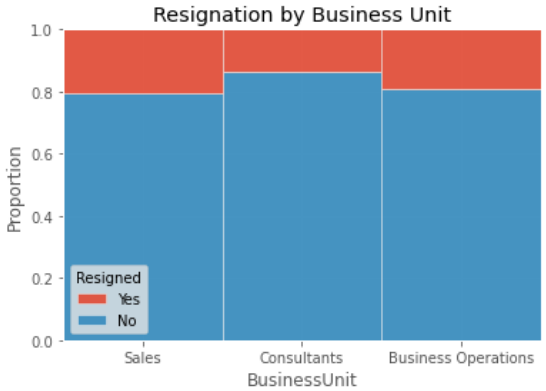


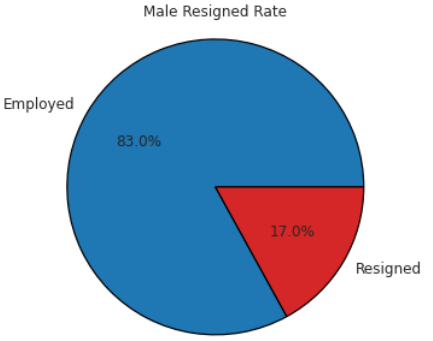




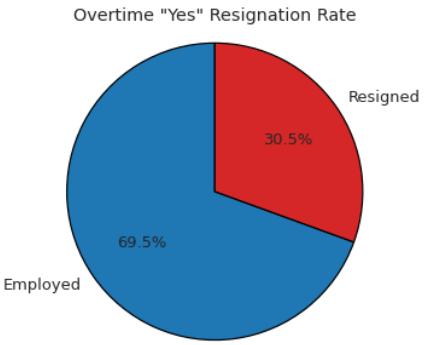


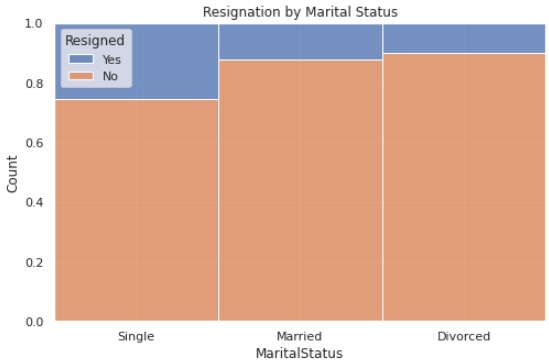


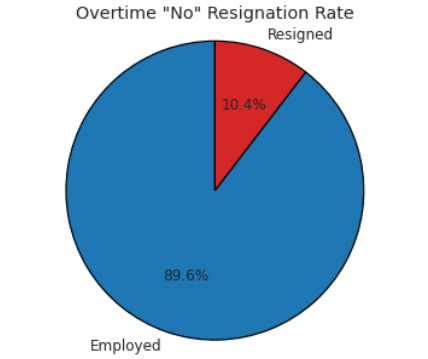




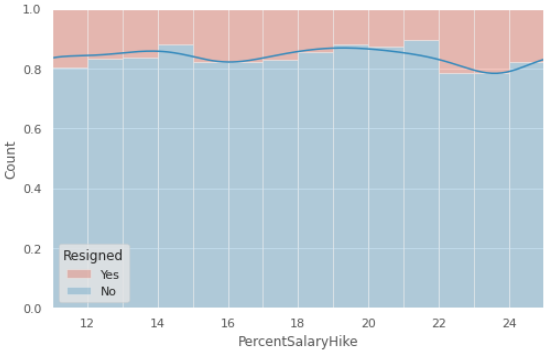

Resignation by Job Satisfaction
Resigned
1.0
1.5
2.0
2.5
3.0
3.5
4.0
JobSatisfaction


Overtime "Yes" Resignation Rate
Resigned
30.5%
69.5%
Employed

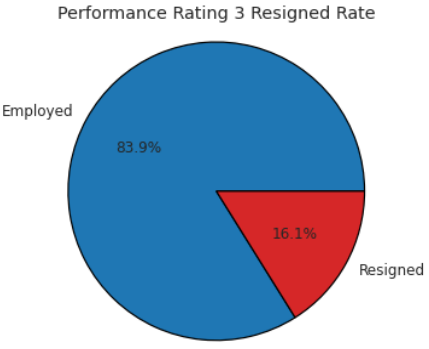

Resignation by Marital Status
Resigned
Yes
No
Count
1.0
0.8
0.6
0.4
0.2
0.0
Single
Married
Divorced
MaritalStatus

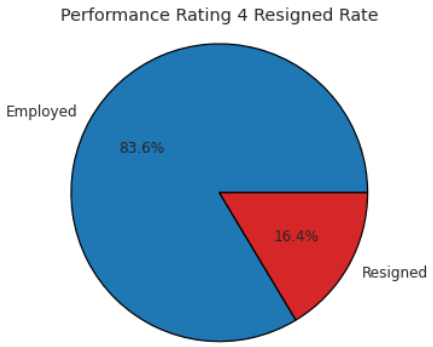

Overtime "No" Resignation Rate
Resigned
10.4%
89.6%
Employed


Resignation by Monthly Income
Yes
No
Resigned
2500
5000
7500
10000
12500
15000
17500
20000
MonthlyIncome

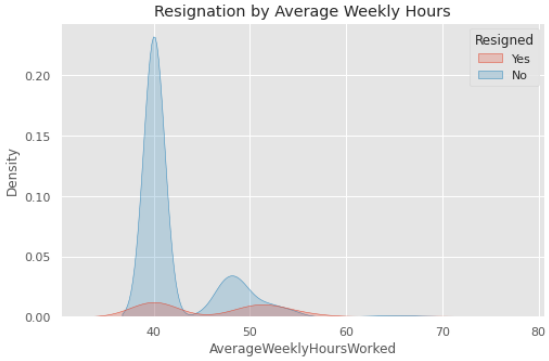

1.0
0.8
0.6
0.4
0.2
0.0
Count
Resigned
Yes
No
12
14
16
18
20
22
24
PercentSalaryHike

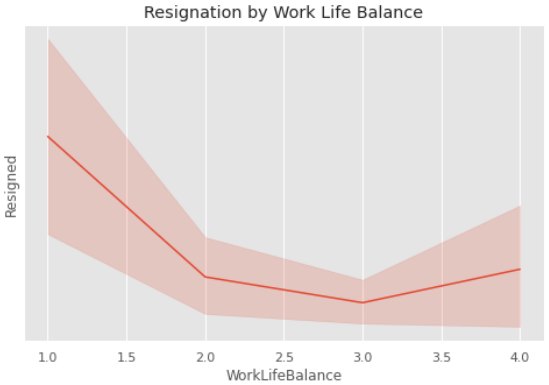

Resignation by Companies Worked
Resigned
Yes
No
count
400
350
300
250
200
150
100
50
0
0
1
2
3
4
5
6
7
8
9
NumCompaniesWorked

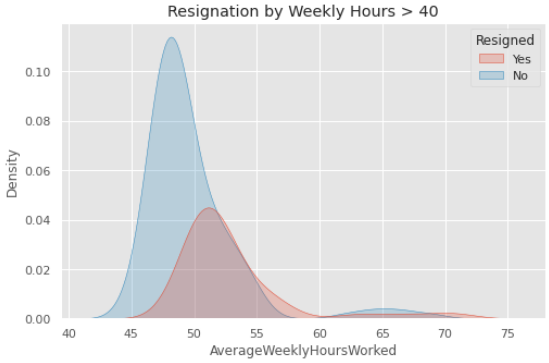

Performance Rating 3 Resigned Rate
Employed
83.9%
16.1%
Resigned

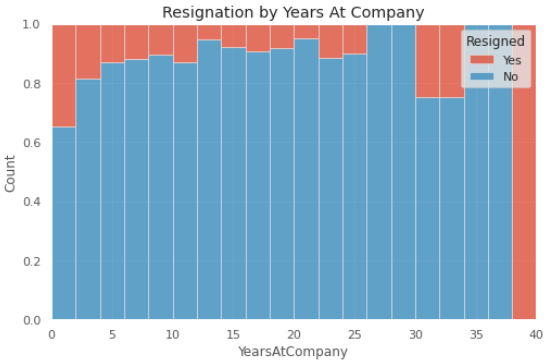

Performance Rating 4 Resigned Rate
Employed
83.6%
16.4%
Resigned

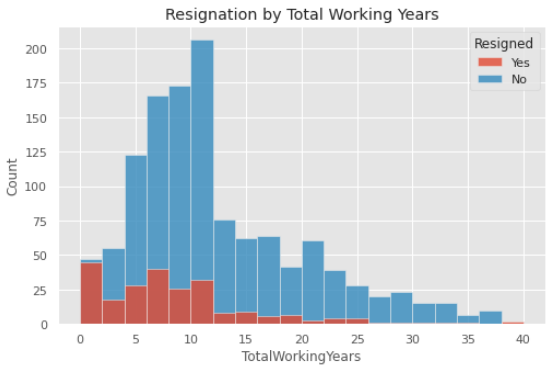

Resignation by Training Last Year
Resigned
Yes
No
count
TrainingTimesLastYear

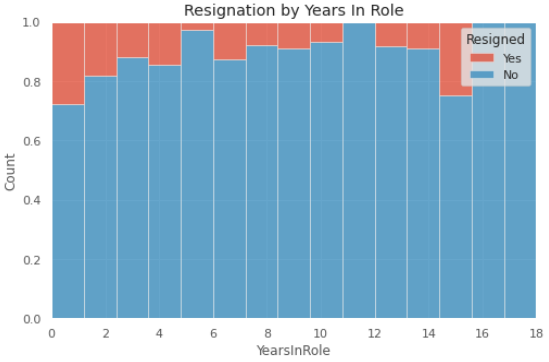

Resignation by Average Weekly Hours
Resigned
Yes
No
Density
AverageWeeklyHoursWorked

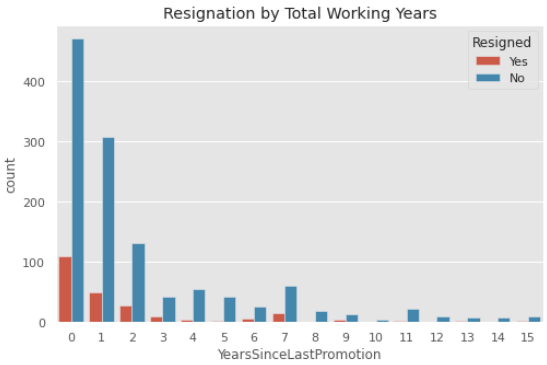

Resignation by Work Life Balance
Resigned
WorkLifeBalance

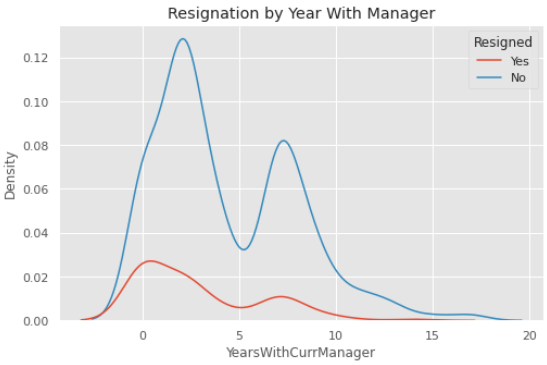

Resignation by Weekly Hours > 40
Resigned
Yes
No
Density
AverageWeeklyHoursWorked

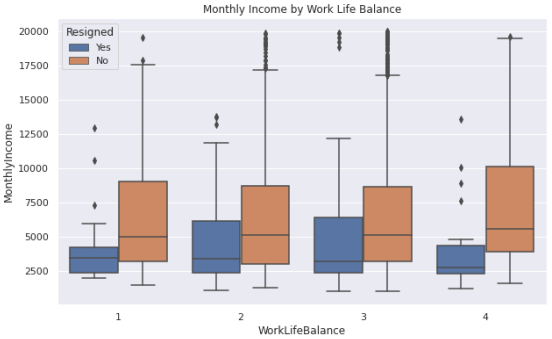

Resignation by Years At Company
Resigned
Yes
No
Count
YearsAtCompany


Resignation by Total Working Years
Resigned
Yes
No
Count
TotalWorkingYears

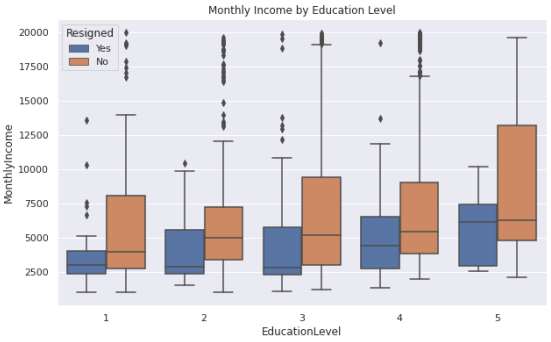

Resignation by Years In Role
Resigned
Yes
No
Count
YearsInRole

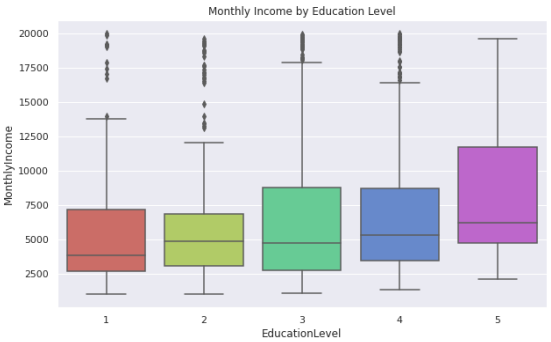


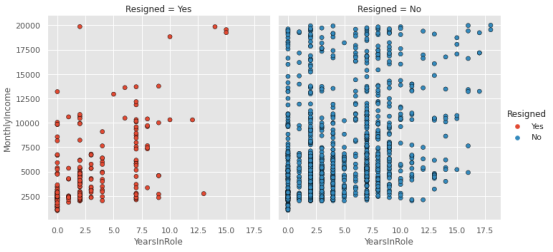


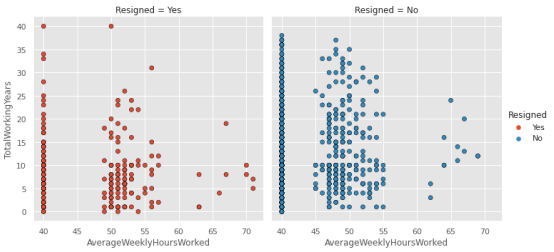


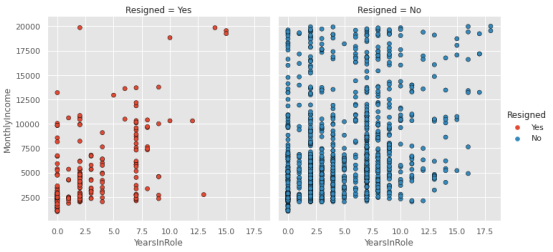
Appendix 2: Relationship Analysis

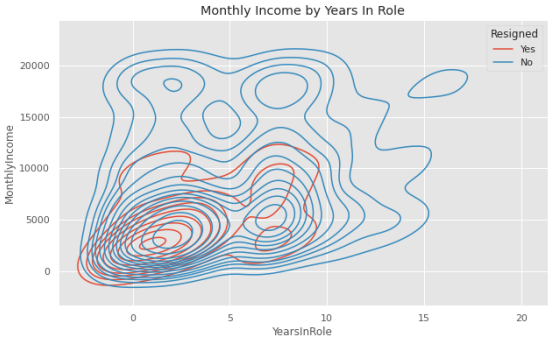






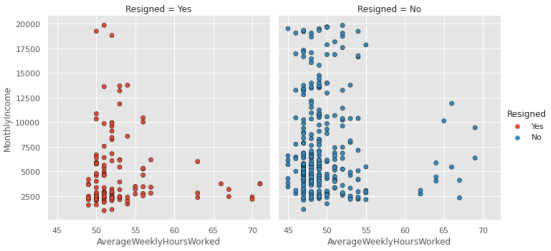


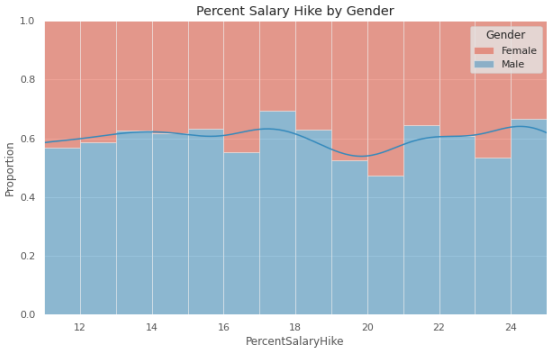



Monthly Income by Years In Role
Resigned
Yes
No
MonthlyIncome
YearsInRole

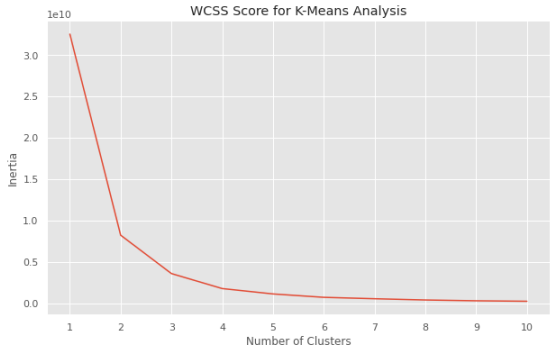

Resigned = Yes
Resigned = No
MonthlyIncome
AverageWeeklyHoursWorked
Resigned
Yes
No

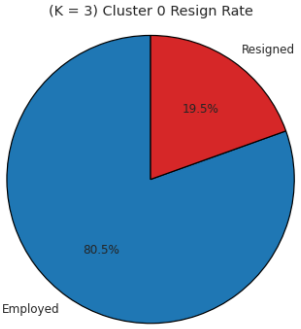

Resigned = Yes
Resigned = No
YearsInRole
AverageWeeklyHoursWorked
Resigned
Yes
No

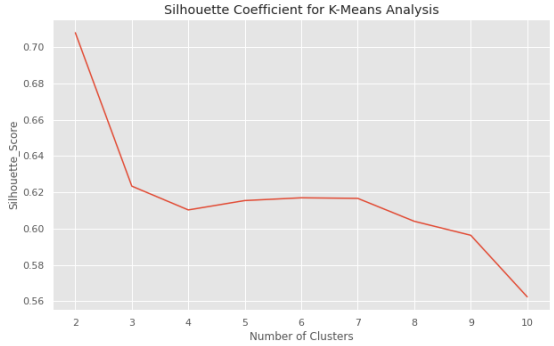

Resigned = Yes
Resigned = No
MonthlyIncome
AverageWeeklyHoursWorked
Resigned
Yes
No

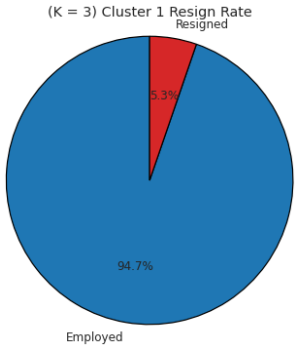

Percent Salary Hike by Gender
Gender
Female
Male
Proportion
PercentSalaryHike

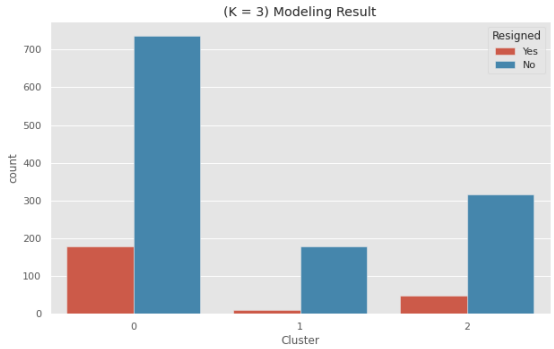
Appendix 3: Modelling Analysis

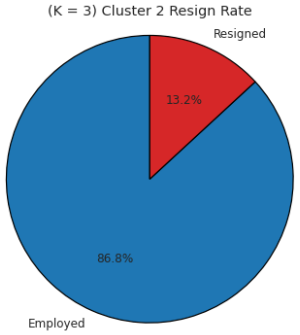


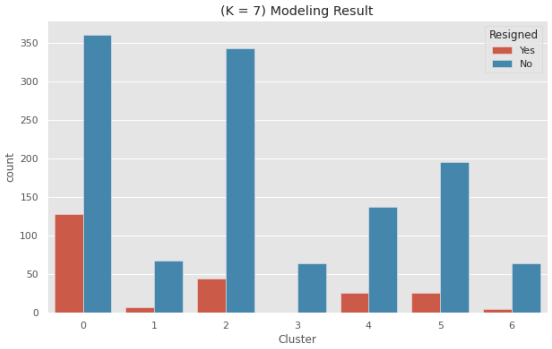


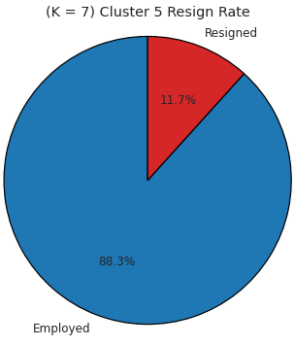


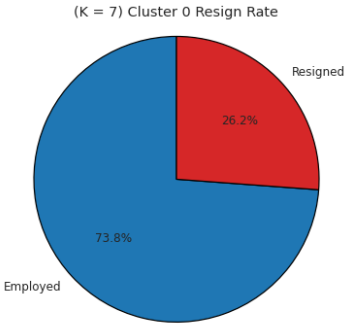


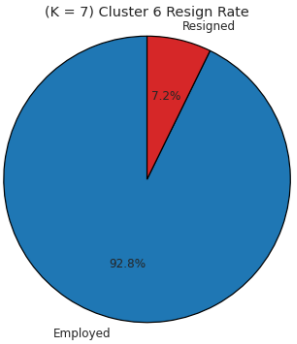


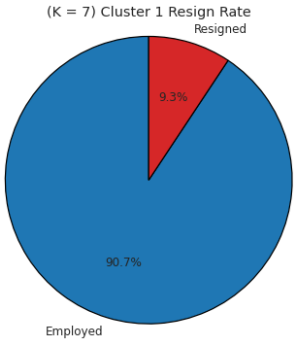


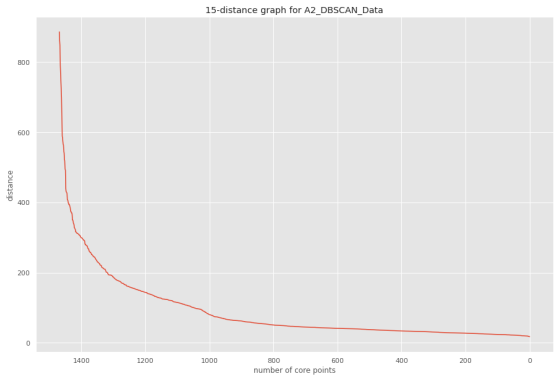

(K = 7) Modeling Result
Resigned
Yes
No
count
Cluster

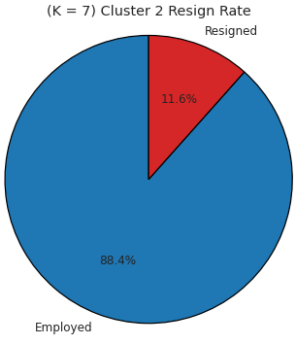

(K = 7) Cluster 5 Resign Rate
Resigned
11.7%
88.3%
Employed

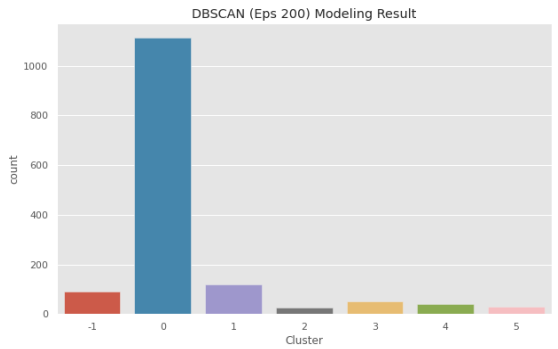

(K = 7) Cluster 0 Resign Rate
Resigned
26.2%
73.8%
Employed

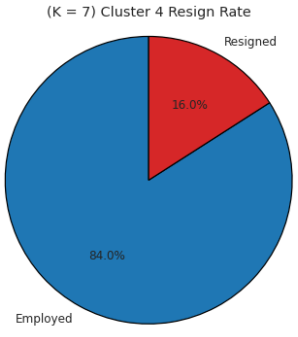

(K = 7) Cluster 6 Resign Rate
Resigned
7.2%
92.8%
Employed

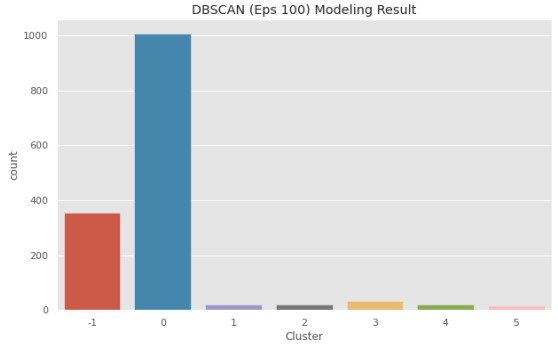

(K = 7) Cluster 1 Resign Rate
Resigned
9.3%
90.7%
Employed

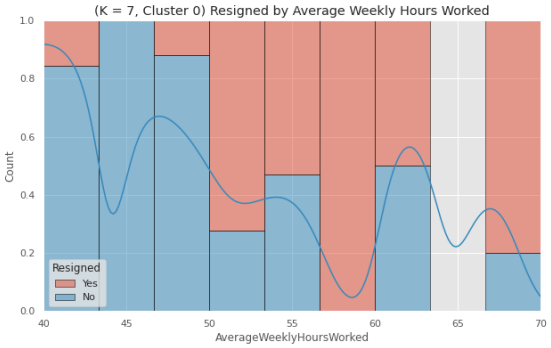

15-distance graph for A2_DBSCAN_Data
distance
number of core points

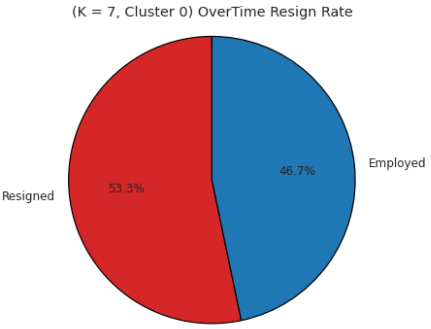

(K = 7) Cluster 2 Resign Rate
Resigned
11.6%
88.4%
Employed

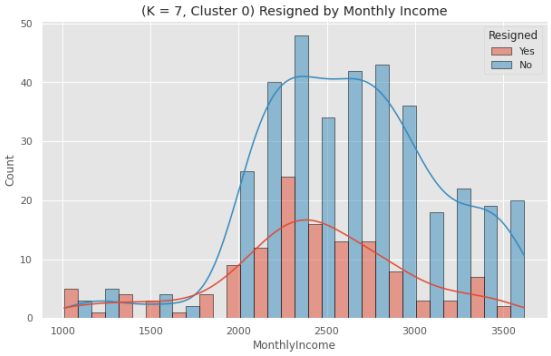

DBSCAN (Eps 200) Modeling Result
count
Cluster


(K = 7) Cluster 4 Resign Rate
Resigned
16.0%
84.0%
Employed


DBSCAN (Eps 100) Modeling Result
count
Cluster


(K = 7, Cluster 0) Resigned by Average Weekly Hours Worked
Count
AverageWeeklyHoursWorked
Resigned
Yes
No


(K = 7, Cluster 0) OverTime Resign Rate
Resigned
53.3%
46.7%
Employed


(K = 7, Cluster 0) Resigned by Monthly Income
Count
MonthlyIncome
Resigned
Yes
No