BIKE MS ANALYTICS

TEAM NAME: KDD DEEP LEARNERS UNCC

Analysis Focus:

"What occupations were responsible for most of BIKE MS's fundraising?"

Motivation:

To identify occupations where people are more likely to contribute towards Multiple Sclerosis and the trend in involvement over the past years.

UNC CHARLOTTE

DATASET USED:

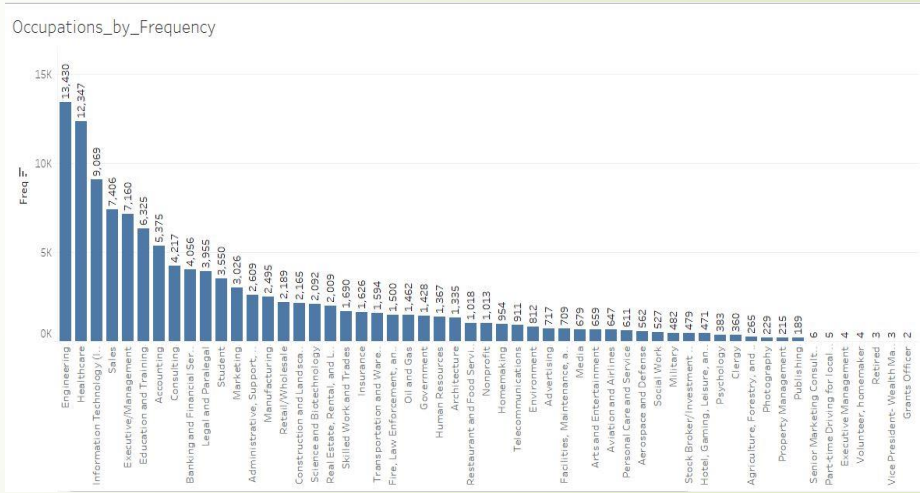
2013-2017 Bike MS Participants AND 2013-2017 Bike Teams data

Exploratory Data Analysis:

#	Parameter	Analysis Technique	Tools Used
1	Email Id	Distribution	R
2	Event Type	Frequency Distribution	R
3	Occupation	Frequency Distribution	R
4	Team Division	Binning by Clustering	R
5	Contribution by Occupation	Time Series Analysis	Tableau, R,

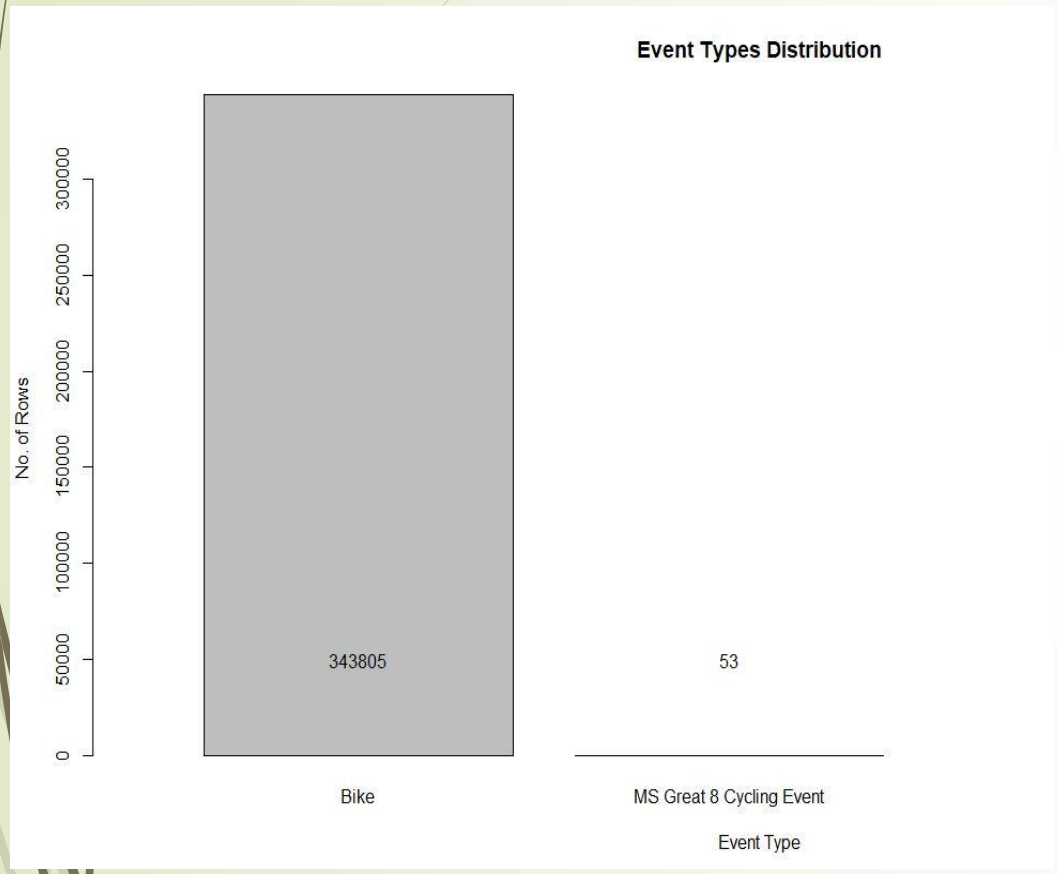
Total Revenue generated by Occupations in 5y sorted by Top 15 Occupations

	Category	x
1	Engineering	14395885.19
2	Executive/Management	12609392.65
3	Healthcare	9495687.64
4	Information Technology (IT)	7606089.71
5	Sales	7597209.04
6	Legal and Paralegal	6634729.17
7	Accounting	5695448.77
8	Consulting	5411887.07
9	Education and Training	4973761.93
10	Banking and Financial Services	4820530.99
11	Marketing	3245598.68
12	Real Estate, Rental, and Leasing	2971337.83
13	Science and Biotechnology	2184393.12
14	Manufacturing	2165160.52
15	Construction and Landscaping	2096496.36



Exploratory Data Analysis

Event Type



MS Great 8 event participation numbers are minimal when compared to Bike MS

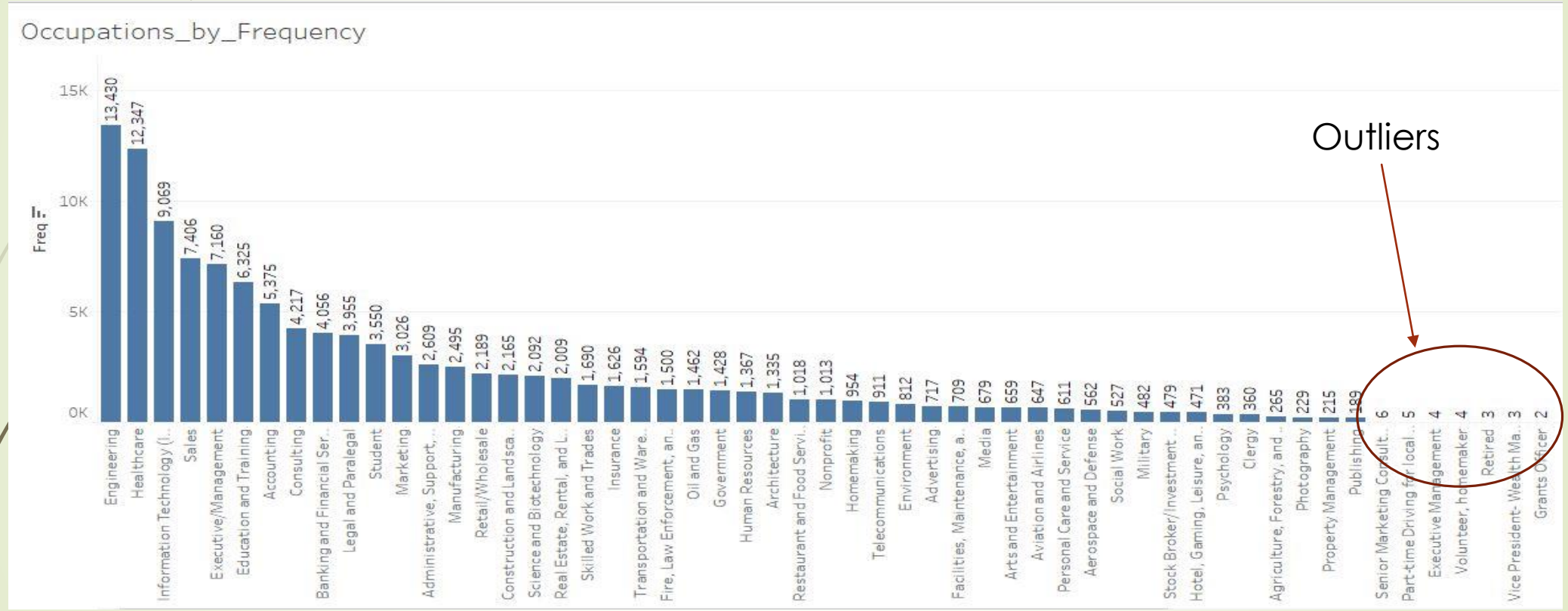
Team Division

	x		
Friends and Family	134698	Organization	92
Corporate	134657	Small Business	90
	34813	Place of worship	85
Family and Friends	14603	Bike Club	81
Corporation	12943	Ohana and Friends	77
Other	3548	Organization (Clubs, Civic Gr	68
Friend and Family	2711	Open	57
Organization (Clubs; Civic Groups; etc.)	2268	Open Team	53
Family/Friends	1320	Ohana	47
Place of Worship	450	Bike Shops	40
Beer/Brewery	411	Organization (Clubs; Civic Gr	25
Bike Shop	305	School	25
School	272	Civic Team	3
Club/Organization	110	Association	2
		Frien`s and Family	2
		Religious	2

Most of the teams came from reference through the friends and family compared to other fields.

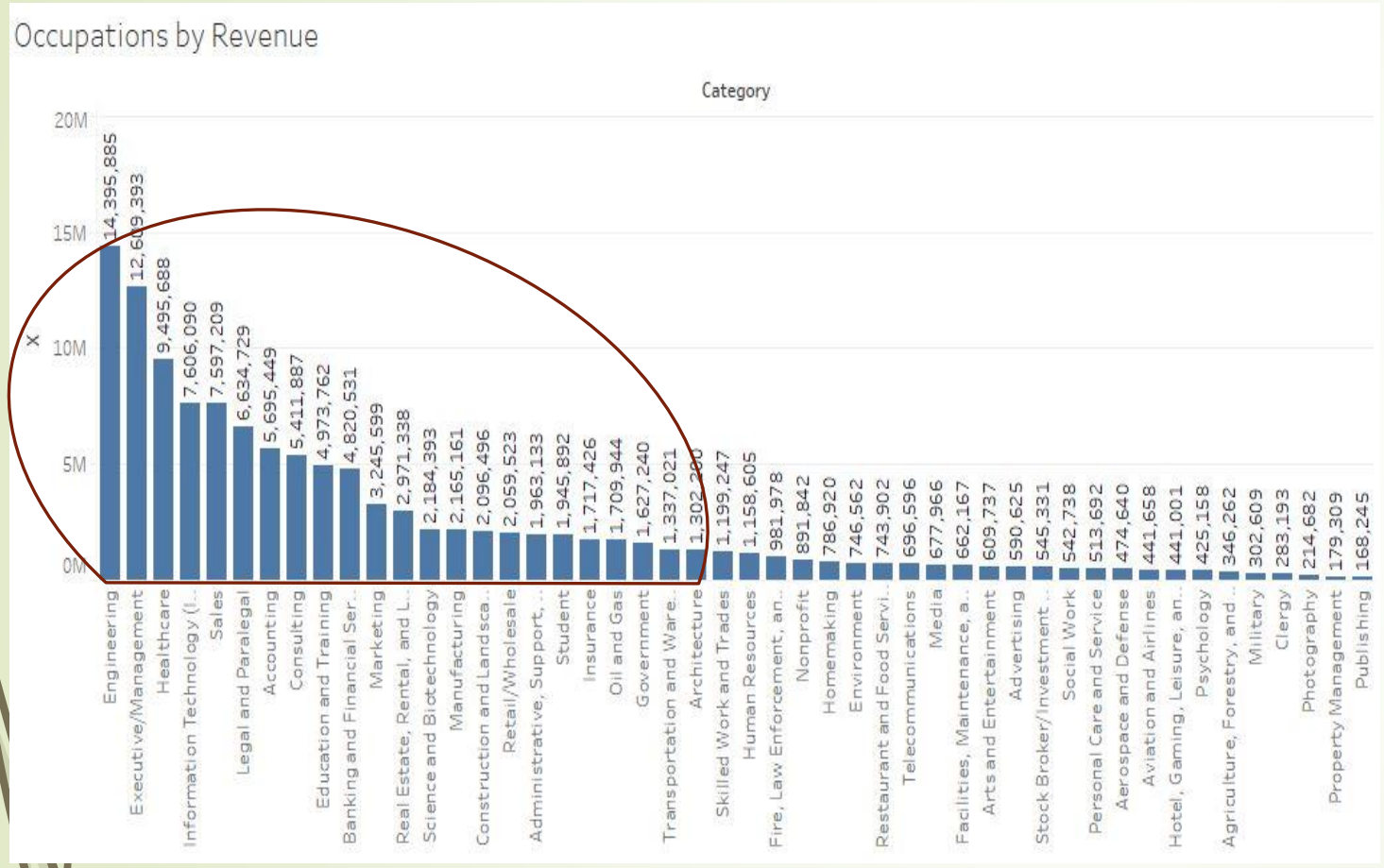
Exploratory Data Analysis

Occupation Distribution



Occupation with frequency less than 10 are omitted considering as outliers.

Exploratory Data Analysis - Occupation Distribution by revenue

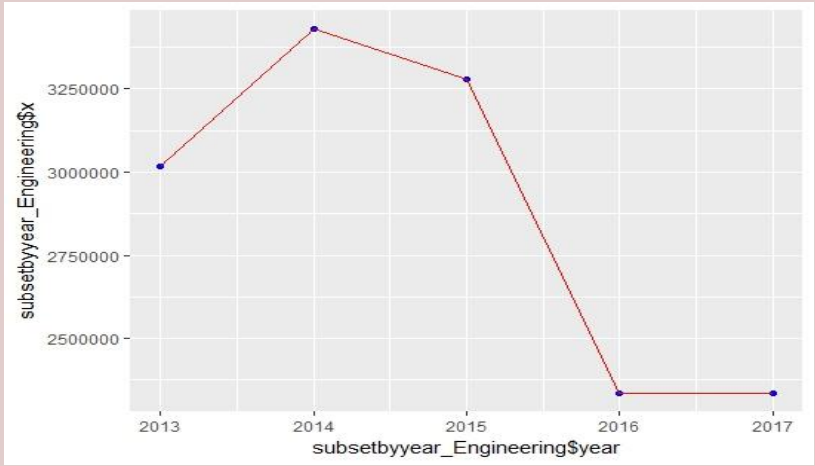
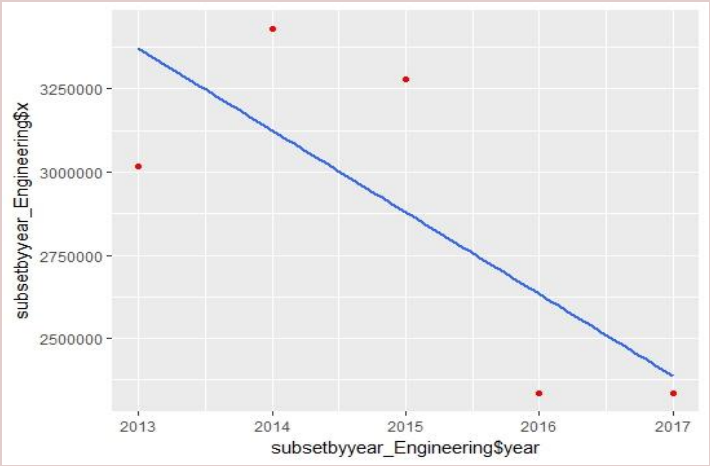
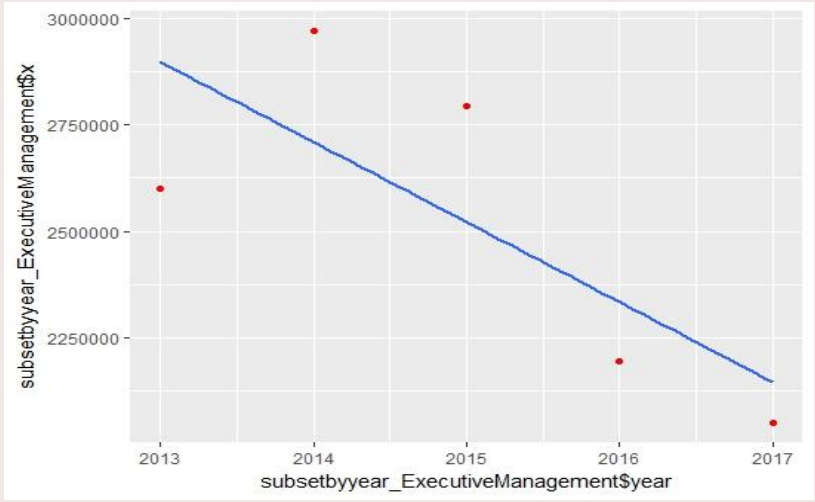
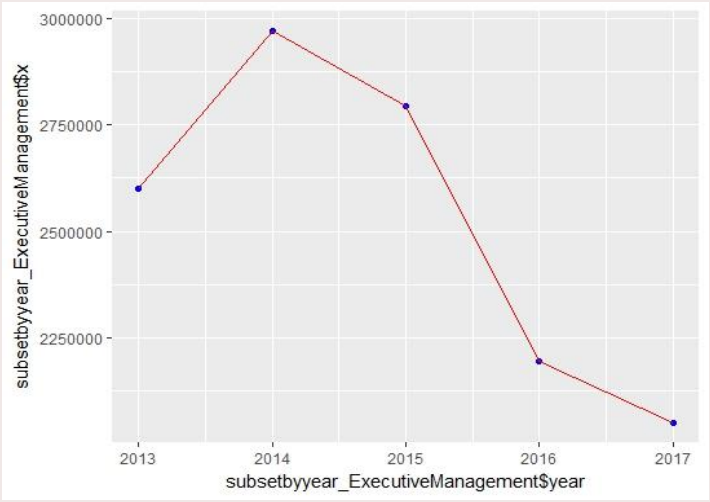


Occupation	2014 Rank	2015 Rank	2016 Rank	2017 Rank
Engineering	1	1	1	1
Executive/Management	2	2	2	2
Healthcare	3	3	3	3
Legal and Paralegal	4	6	6	6
Information Technology (IT)	5	4	5	5
Sales	6	5	4	4
Accounting	7	7	7	7
Consulting	8	8	8	8
Education and Training	9	9	9	9
Banking and Financial Services	10	10	10	10

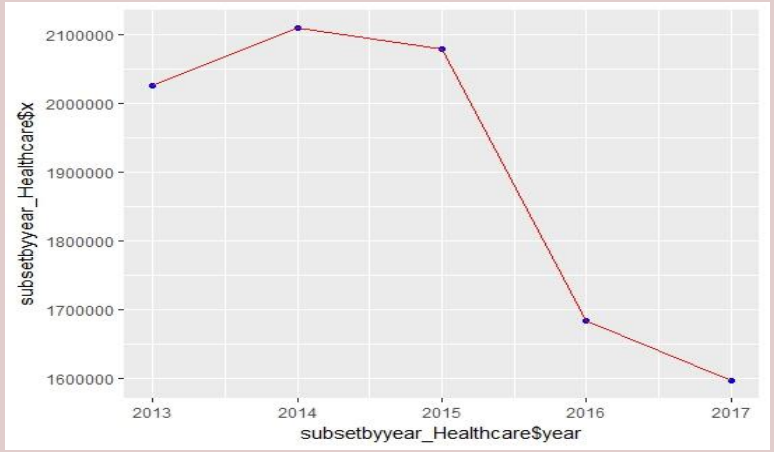
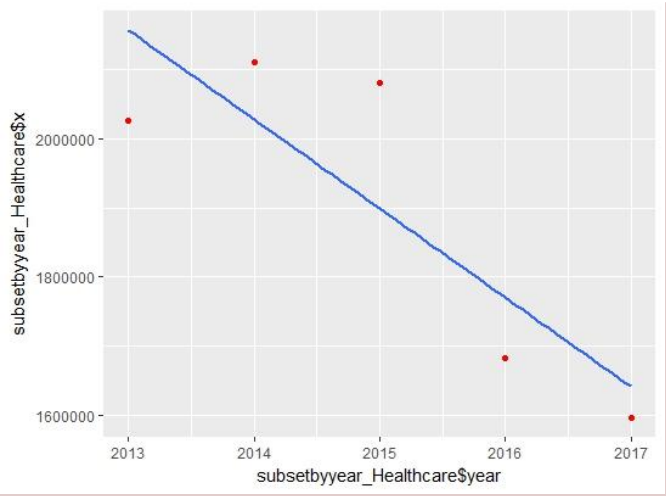
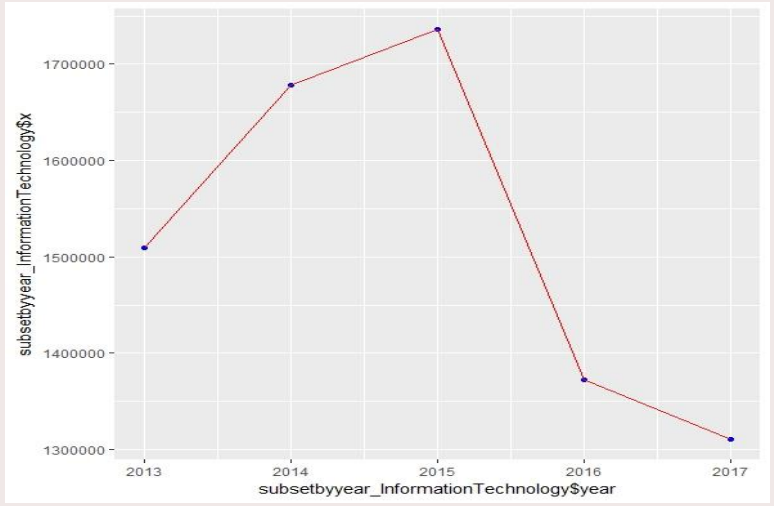
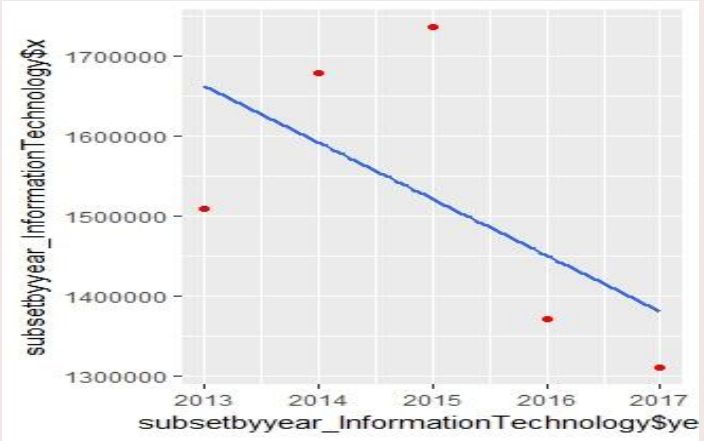
The graph shows the top performing occupations by revenue.

Contributions by year
Sorted by top 5 occupations

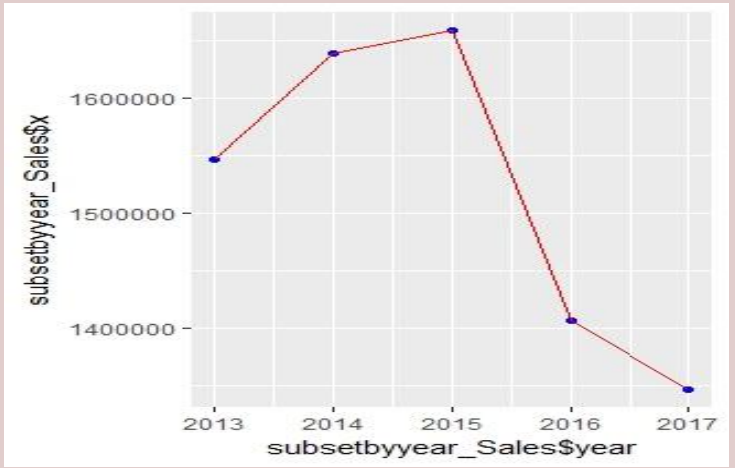
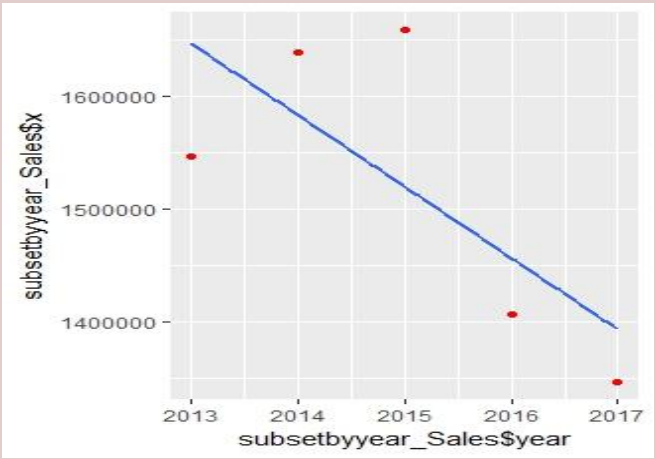
MODELLING – TIME SERIES ANALYSIS

#	Occupation	Aggregated by year map	Linear regression
1	Engineering		
2	Executive/Management		

MODELLING – TIME SERIES ANALYSIS

#	Occupation	Aggregated by year map	Linear regression
3	Health Care		
4	Information Technology		

MODELLING – TIME SERIES ANALYSIS

#	Occupation	Aggregated by year map	Linear regression
5	Sales		

PREDICTION RESUTLS

Occupation	Linear regression equation	Year	Predicted Amount
Engineering	Amount = 245701*(Year) - 497966648	2018	2142030
Executive/management	Amount = 187709*(Year) - 380756779	2018	1960017
Health care	Amount = 128556 *(Year) - 260940996	2018	1514988
Information technology	Amount = 70309 *(Year) - 143194401	2018	1310839
Sales	Amount = 63189 *(Year) - 128845442	2018	1330040