

**ANL252 (Online)**

**PYTHON FOR DATA ANALYTICS**

# **Tutor-Marked Assignment**

**July 2022 Presentation**

**Submitted by:**

|  |  |
| --- | --- |
| **Name** | **PI No.** |
| **Toh Li Ting** | **M2172239** |

**Submission Date: 14/08/2022**

**Question 1**

**a)**

Chart 1:

Above bar chart, is a comparison of the Satisfaction level of the total 250 former and current employee by gender. Having the orange bar as the female and the blue bar as the male reading.

|  |  |  |  |
| --- | --- | --- | --- |
| Satisfaction Score | |  |  |
| *Frequency* | *Bin* | *F* | *M* |
| 1 | 1 | 1 | 0 |
| 4 | 2 | 1 | 3 |
| 84 | 3 | 45 | 39 |
| 80 | 4 | 48 | 32 |
| 81 | 5 | 49 | 32 |
| 250 | Total | 144 | 106 |

From the chart we are able to see that the company have a total 144 female employees and 106 male employees in the collected data. With the mode of the satisfaction score at 5 for female and 3 for male employees and the mean of 3.99 for female and 3.88 for male employee. Which we are able to say that the females are slightly more satisfied with their job.

|  |  |  |
| --- | --- | --- |
|  | *F* | *M* |
| mean | 3.99 | 3.88 |
| median | 4 | 4 |
| mode | 5 | 3 |
|  |  |  |

Chart 2:

Chart 2 is the comparison of the Employment Status by Marital Status. With orange being the current employees and blue being the former employees.

From the chart we are able to see that there are a higher ratio of the employee leaving the company if they are divorcee with 50% left the company through the years, and 40% for married employees, 27% for singles and 19% for others. Therefore, we are able to say that singles are last likely to leave their job.

**b)**

Chart 3:



# importing the required module

import matplotlib.pyplot as plt

# x axis values

x = [1,2,3,4,5]

# corresponding y axis values

y = [1,1,45,48,49]

# plotting the points

plt.plot(x, y)

# naming the x axis

plt.xlabel('Satisfaction Score')

# naming the y axis

plt.ylabel('No. of Female Employee')

# giving a title to my graph

plt.title('Employee Satisifaction Level By Gender (Female)')

# function to show the plot

plt.show()

Employment Status by Marital Status

Chart 4:

# importing the required module

import matplotlib.pyplot as plt

# x axis values

x = [1,2,3,4,5]

# corresponding y axis values

y = [0,3,39,32,32]

# plotting the points

plt.plot(x, y)

# naming the x axis

plt.xlabel('Satisfaction Score')

# naming the y axis

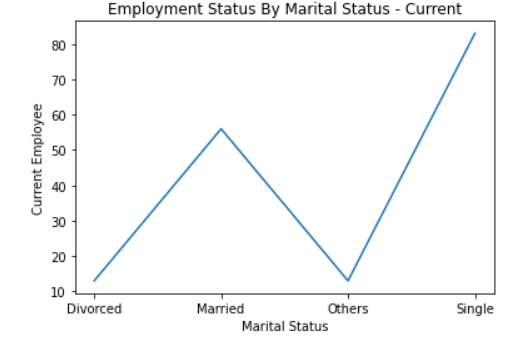
plt.ylabel('No. of Female Employee')

# giving a title to my graph

plt.title('Employee Satisifaction Level By Gender (Male)')

# function to show the plot

plt.show()

Chart 5:

# importing the required module

import matplotlib.pyplot as plt

# x axis values

x = ['Divorced','Married','Others','Single']

# corresponding y axis values

y = [13,56,13,83]

# plotting the points

plt.plot(x, y)

# naming the axises

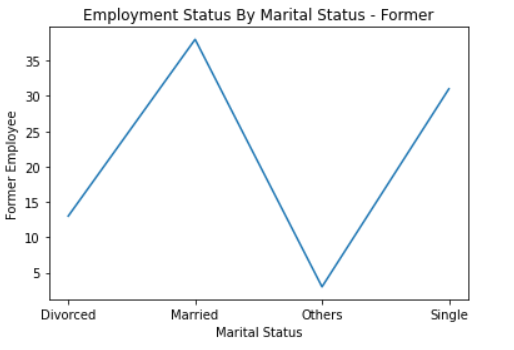
plt.xlabel('Marital Status')

plt.ylabel('Current Employee')

plt.title('Employment Status By Marital Status - Current')

# function to show the plot

plt.show()

Chart 6:

# importing the required module

import matplotlib.pyplot as plt

# x axis values

x = ['Divorced','Married','Others','Single']

# corresponding y axis values

y = [13,38,3,31]

# plotting the points

plt.plot(x, y)

# naming the x axis

plt.xlabel('Marital Status')

# naming the y axis

plt.ylabel('Former Employee')

# giving a title to my graph

plt.title('Employment Status By Marital Status - Former')

# function to show the plot

plt.show()

**c)**

Staff Employment JoinDate LeftDate

0 Aairah Hastings Employee 2012-04-02 2022-05-01

1 Aaminah Mcmillan Employee 2014-05-12 2022-05-01

2 Aamna Howell Employee 2014-09-29 2022-05-01

3 Aayan Ferrell Employee 2016-01-05 2022-05-01

4 Ace Potter Employee 2011-07-11 2020-09-06

.. ... ... ... ...

245 Zaina Parrish Employee 2015-02-16 2022-05-01

246 Zane Downes Employee 2017-04-20 2022-05-01

247 Zayd Farley Employee 2014-01-05 2022-05-01

248 Zoya Easton Employee 2012-08-13 2015-09-01

249 Zuzanna Pruitt Employee 2015-03-30 2022-05-01

[250 rows x 4 columns]

import numpy as np

import pandas as pd

import datetime as datetime

data = pd.DataFrame({'p1':Param1,'p2': Param2,’})

startDT = data['p1']

endDT = data['p2']

dateDiff = pd.to\_datetime(endDT, errors='coerce') - pd.to\_datetime(startDT, errors='coerce')

outputDF = pd.DataFrame({'Start Date': startDT,'End Date': endDT,'Difference (Days)': dateDiff})

**d)**

name = ("Aairah Hastings","Aaminah Mcmillan","Aamna Howell")

employment\_status = ("employed")

while True:

user\_input = input("What is the name of the employee?")

if user\_input in name:

name\_index = name.index(user\_input)

employment\_status = employment\_status

print(f"The staff of name {user\_input} is {employment\_status} by the company.")

break

else:

user\_input = input("Invalid input. Please try again.")