|  |  |
| --- | --- |
| Project Title | IBM HR Analytics Employee Attrition & Performance |
| Tools | Python, ML, SQL, Excel |
| Domain | Data Analyst & Data scientist |
| Project Difficulties level | intermediate |

Dataset : Dataset is available in the given link. You can download it at your convenience.

[Click](https://drive.google.com/file/d/1tv2dfgghMKzN_mKmdn0ymCUrkwVvBk3w/view?usp=sharing) [here](https://drive.google.com/file/d/1tv2dfgghMKzN_mKmdn0ymCUrkwVvBk3w/view?usp=sharing) [to](https://drive.google.com/file/d/1tv2dfgghMKzN_mKmdn0ymCUrkwVvBk3w/view?usp=sharing) [download](https://drive.google.com/file/d/1tv2dfgghMKzN_mKmdn0ymCUrkwVvBk3w/view?usp=sharing) [data](https://drive.google.com/file/d/1tv2dfgghMKzN_mKmdn0ymCUrkwVvBk3w/view?usp=sharing) [set](https://drive.google.com/file/d/1tv2dfgghMKzN_mKmdn0ymCUrkwVvBk3w/view?usp=sharing)

About Dataset

Uncover the factors that lead to employee attrition and explore important questions such as ‘show me a breakdown of distance from home by job role and attrition’ or ‘compare average monthly income by education and attrition’. This is a fictional data set created by IBM data scientists.

Education

1. 'Below College'
2. 'College'
3. 'Bachelor'
4. 'Master'
5. 'Doctor'

EnvironmentSatisfaction

1. 'Low'
2. 'Medium'
3. 'High'
4. 'Very High'

JobInvolvement

1. 'Low'
2. 'Medium'
3. 'High'
4. 'Very High'

JobSatisfaction

1. 'Low'
2. 'Medium'
3. 'High'
4. 'Very High'

PerformanceRating

1. 'Low'
2. 'Good'
3. 'Excellent'
4. 'Outstanding'

RelationshipSatisfaction

1. 'Low'
2. 'Medium'
3. 'High'
4. 'Very High' WorkLifeBalance
5. 'Bad'
6. 'Good'
7. 'Better'
8. 'Best'

**Example: You can get the basic idea how you can create a project from here** what steps you should have to follow

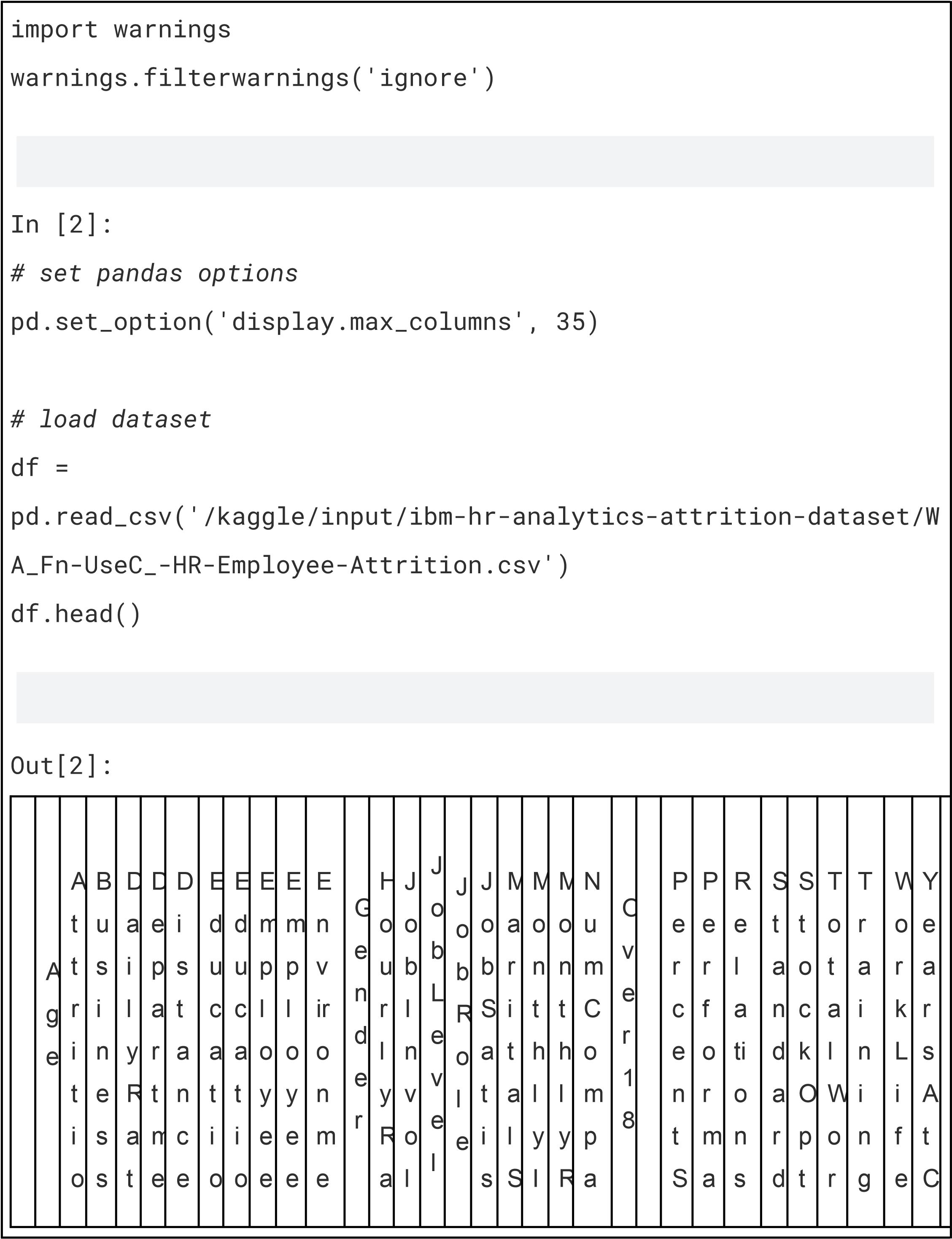
**Example: You can get the basic idea how you can create a project from here**

Sample code and output

HR Attrition Analysis[¶](https://www.kaggle.com/code/febbyanggraini/hr-attrition-eda#HR-Attrition-Analysis)

In the business world, companies often face the challenge of retaining talented employees. One of the most pressing issues is the increasing rate of employee turnover, commonly known as HR attrition. Turnover can have a significant impact on a company’s productivity, stability, and long-term sustainability. High attrition rates

|  |
| --- |
| can lead to increased recruitment and training costs, disrupt team dynamics, and result in the loss of valuable institutional knowledge. Therefore, understanding the factors contributing to attrition and implementing effective retention strategies is crucial for maintaining a competitive edge and ensuring  Objectives of the Analysis   1. Understand Current Turnover Rates: Gain a comprehensive understanding of the current employee turnover rate and analyze the demographic distribution of attrition by age, gender, education, department, and job role. 2. Identify Key Factors Influencing Turnover: Examine the main factors contributing to employee attrition, including job satisfaction indicators (job involvement and work-life balance), salary factors (monthly income and salary hikes), and benefit factors (stock option levels), to uncover patterns and correlations that drive higher attrition rates.   Data Cleaning  In [1]:  *# import data manipulation package* import pandas as pd import numpy as np  *# import data visualization package* import matplotlib.pyplot as plt import seaborn as sns  *# importing the warnings library* |

O v e r

T i m e

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  |  | n | T  r  a v e  l | e | n  t | F  r  o  m H  o  m  e | n | n F  i  e  l  d | C  o u n  t | N  u  m  b e  r | n  t  S a  ti s  f  a c ti o n |  | t  e | v e  m  e n  t |  |  | f  a c  t i  o n | t  a  t  u s | n c o  m  e | a  t  e | n  i  e s  W  o  r  k e d |  |  | a  l  a  r  y  H  i  k e | n c e R  a  t i  n g | h  i  p S a  ti s  f  a c ti o n | H  o u  r  s | i  o n L e v e  l | k  i  n g Y e a  r  s | T  i  m  e s L a s  t  Y e a  r | B a  l  a n c e | o  m  p a n y |  |
| 0 | 4  1 | Y e s | T  r  a v e  l  \_ R  a  r  e  l | 1 1 0 2 | S a  l  e s | 1 | 2 | L  i f  e S  c  i  e n c e s | 1 | 1 | 2 | F e  m  a  l  e | 9  4 | 3 | 2 | S a  l  e s  E  x e c u  t i  v | 4 | S  i  n g  l  e | 5 9 9 3 | 1 9 4 7 9 | 8 | Y | Y e s | 1  1 | 3 | 1 | 8  0 | 0 | 8 | 0 | 1 | 6 |  |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  |  |  | y |  |  |  |  |  |  |  |  |  |  |  |  | e |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 4  9 | N  o | T r  a v e l  \_ F r  e q u e n  t l  y | 2  7 9 | R  e s e a  r  c h  & D  e v e  l  o p  m  e n  t | 8 | 1 | L  i f  e S  c  i  e n c e s | 1 | 2 | 3 | M  a  l  e | 6  1 | 2 | 2 | R e s e a r  c h S c  i  e n  t i  s  t | 2 | M a r r i  e d | 5 1 3 0 | 2  4  9  0 7 | 1 | Y | N  o | 2  3 | 4 | 4 | 8  0 | 1 | 1  0 | 3 | 3 | 1  0 |  |
| 2 | 3  7 | Y e | T  r  a | 1 3 7 | R  e s | 2 | 2 | O  t  h | 1 | 4 | 4 | M  a  l | 9  2 | 2 | 1 | L a b | 3 | S  i  n | 2 0 9 | 2 3 9 | 6 | Y | Y e | 1  5 | 3 | 2 | 8  0 | 0 | 7 | 3 | 3 | 0 |  |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  |  | s | v e  l  \_ R  a  r  e  l  y | 3 | e a  r  c h  & D  e v e  l  o p  m  e n  t |  |  | e  r |  |  |  | e |  |  |  | o  r  a  t  o  r  y T e c h n  i  c  i  a n |  | g  l  e | 0 | 6 |  |  | s |  |  |  |  |  |  |  |  |  |  |
| 3 | 3  3 | N  o | T  r  a v e  l  \_ F  r | 1 3 9 2 | R  e s e a  r  c h  & | 3 | 4 | L  i f  e S  c  i  e n | 1 | 5 | 4 | F e  m  a  l  e | 5  6 | 3 | 1 | R  e s e a  r  c h  S | 3 | M  a  r r  i  e d | 2 9 0 9 | 2 3 1 5 9 | 1 | Y | Y e s | 1  1 | 3 | 3 | 8  0 | 0 | 8 | 3 | 3 | 8 |  |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  |  |  | e q u e n  t l  y |  | D  e v e  l  o p  m  e n  t |  |  | c e s |  |  |  |  |  |  |  | c  i  e n  t i  s  t |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | 2  7 | N  o | T  r  a v e  l  \_ R  a  r  e  l  y | 5 9 1 | R  e s e a  r  c h  & D  e v e  l  o | 2 | 1 | M  e d  i  c a  l | 1 | 7 | 1 | M  a  l  e | 4  0 | 3 | 1 | L a b o  r  a  t  o  r  y T e c h n | 2 | M  a  r r  i  e d | 3 4 6 8 | 1 6 6 3 2 | 9 | Y | N  o | 1  2 | 3 | 4 | 8  0 | 1 | 6 | 3 | 3 | 2 |  |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

p

m

e

n

t

i

c

i

a

n

In

[3]:

*#*

*check*

*data*

*shape*

df.shape

Out[3]:

(1470

,

35)

The

output

indicates

that

the

DataFrame

has

1,470

rows

and

35

columns.

In

[4]:

*#*

*check*

*number*

*of*

*dupliacted*

*data*

print(f'Number

of

duplicated

data:

df.duplicated().sum()}'

{

)

Number

of

duplicated

data:

0

In

[5]:

*#*

*check*

*missing*

*values*

|  |
| --- |
| df.isnull().sum() / len(df) \* 100  Out[5]:  Age 0.0  Attrition 0.0  BusinessTravel 0.0  DailyRate 0.0  Department 0.0  DistanceFromHome 0.0  Education 0.0  EducationField 0.0  EmployeeCount 0.0  EmployeeNumber 0.0  EnvironmentSatisfaction 0.0  Gender 0.0  HourlyRate 0.0  JobInvolvement 0.0  JobLevel 0.0  JobRole 0.0  JobSatisfaction 0.0  MaritalStatus 0.0  MonthlyIncome 0.0  MonthlyRate 0.0 |

|  |
| --- |
| NumCompaniesWorked 0.0  Over18 0.0  OverTime 0.0  PercentSalaryHike 0.0  PerformanceRating 0.0  RelationshipSatisfaction 0.0  StandardHours 0.0  StockOptionLevel 0.0  TotalWorkingYears 0.0  TrainingTimesLastYear 0.0  WorkLifeBalance 0.0  YearsAtCompany 0.0  YearsInCurrentRole 0.0  YearsSinceLastPromotion 0.0 YearsWithCurrManager 0.0  dtype: float64  There are any missing values in the dataset.  In [6]:  *# check data types* df.dtypes  Out[6]: |

|  |  |
| --- | --- |
| Age | int64 |
| Attrition | object |
| BusinessTravel | object |
| DailyRate | int64 |
| Department | object |
| DistanceFromHome | int64 |
| Education | int64 |
| EducationField | object |
| EmployeeCount | int64 |
| EmployeeNumber | int64 |
| EnvironmentSatisfaction | int64 |
| Gender | object |
| HourlyRate | int64 |
| JobInvolvement | int64 |
| JobLevel | int64 |
| JobRole | object |
| JobSatisfaction | int64 |
| MaritalStatus | object |
| MonthlyIncome | int64 |
| MonthlyRate | int64 |
| NumCompaniesWorked | int64 |
| Over18 | object |
| OverTime | object |
| PercentSalaryHike | int64 |
| PerformanceRating | int64 |

StandardHours

RelationshipSatisfaction

int64

int64

int64

int64

int64

int64

int64

int64

int64

int64

All

columns

have

appropriate

data

types,

ensuring

that

the

data

is

correctly

formatted

A

g

e

D

a

i

l

D

i

s

t

E

d

u

c

E

m

p

l

E

m

p

l

E

n

vi

ro

H

o

u

r

J

o

b

I

J

o

b

L

J

o

b

S

M

o

n

t

M

o

n

t

N

u

m

C

P

e

r

c

P

e

rf

o

R

el

at

io

S

t

a

n

S

t

o

c

T

o

t

a

Tr

ai

ni

n

W

o

r

k

Y

e

a

r

Y

e

a

r

Y

e

ar

s

Y

e

ar

s

StockOptionLevel

TotalWorkingYears

TrainingTimesLastYear

WorkLifeBalance

YearsAtCompany

YearsInCurrentRole

YearsSinceLastPromotion YearsWithCurrManager dtype: object

for analysis.

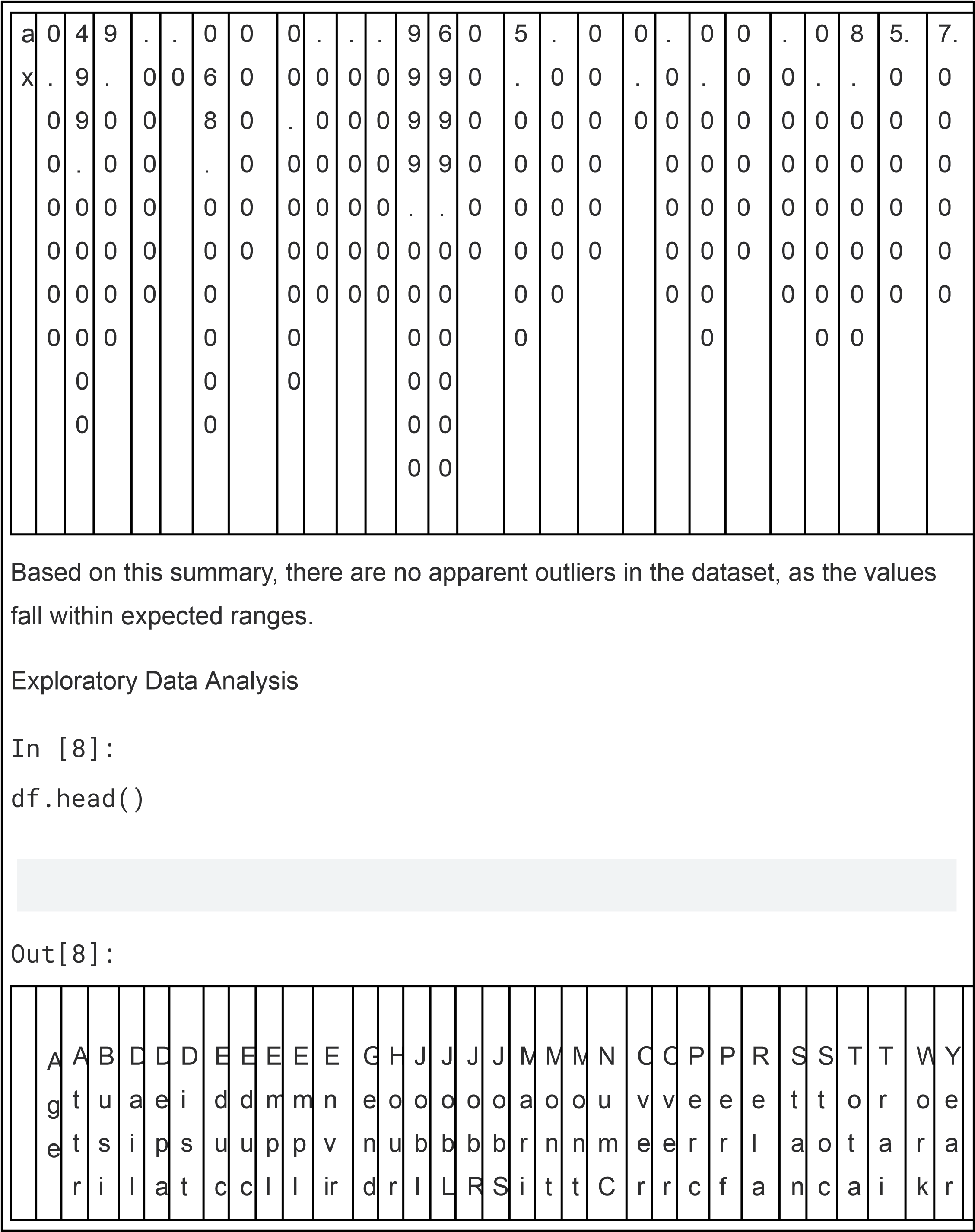
In [7]: *# check data decribe* df.describe() Out[7]:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  |  | y  R  a  t  e | a n c e F r  o  m H o  m  e | a t i  o n | o y e e C o u n  t | o y e e N u  m  b e  r | n  m  e  nt  S  at  isf  a  cti o n | l  y  R  a t  e | n v o  l  v e  m  e n  t | e v e  l | a t i  s f  a c  t i  o n | h l  y  I  n c o  m  e | h l  y  R a t  e | o  m  p a  ni e s  W or k e d | e n  t  S a l  a  r  y  H  i  k e | r  m  a n c e R  a ti n g | n s  hi  p S  at is fa ct io  n | d a  r  d H o u  r  s | k  O  p  ti o n L e v e  l | l  W  o  r  k i  n g Y e a  r  s | g  Ti m  e s L a  st  Y e  ar | L i f  e B a  l  a n c e | s  A  t  C o  m  p a n y | s  I  n C u  rr  e n t  R  o  l  e | Si nc e L as tP ro  m  oti o n | W it h C  ur r  M  a n a g  er |
| c o u n  t | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0.  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0.  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0.  0 0 0 0 0 0 | 1 4 7  0  .  0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0.  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0  .  0 0 0 0 0 0 | 1 4 7  0.  0 0 0 0 0 0 | 1 4 7  0.  0 0 0 0 0 0 |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | m e a n | 3  6  .  9  2  3  8 1 0 | 8 0  2  .  4 8 5  7  1 4 | 9  .  1  9 2 5  1 7 | 2  .  9   1. 2 9 2. 5 | 1  .  0 | 1  0  2  4  .  8  6  5 3 0 6 | 2.  7  2  1  7  6 9 | 6  5  .  8  9  1  1 5 6 | 2  .  7   1. 9 9 2. 2 | 2  .  0  6 3 9  4 6 | 2  .  7  2 8 5  7 1 | 6  5  0  2  .  9  3  1 2 9 3 | 1  4 3 1  3  .  1 0  3  4 0 1 | 2.  6  9  3  1  9 7 | 1  5  .  2  0  9  5 2 4 | 3  .  1  5 3 7  4 1 | 2.  7  1  2  2  4 5 | 8  0  .  0 | 0  .  7  9 3 8  7 8 | 1  1  .  2  7  9  5 9 2 | 2.  7  9  9  3  2 0 | 2  .  7  6 1 2  2 4 | 7  .  0  0 8 1  6 3 | 4  .  2  2 9 2  5 2 | 2.  1  8  7  7  5 5 | 4.  1  2  3  1  2 9 |
| s  t  d | 9  .  1 3 5 3 7 3 | 1. 0   3  .   1. 0 9 1 0 0 | 8  .  1 0 6 8 6 4 | 1  .  0 2 4 1 6 5 | 0  .  0 | 6 0  2  .  0 2 4 3 3 5 | 1.  0 9 3 0 8 2 | 2  0  .  3 2 9 4 2 8 | 0  .  7 1 1 5 6 1 | 1  .  1 0 6 9 4 0 | 1  .  1 0 2 8 4 6 | 4 7 0  7  .  9 5 6 7 8 3 | 7 1 1  7  .  7 8 6 0 4 4 | 2.  4 9 8 0 0 9 | 3  .  6 5 9 9 3 8 | 0  .  3 6 0 8 2 4 | 1.  0 8 1 2 0 9 | 0  .  0 | 0  .  8 5 2 0 7 7 | 7  .  7 8 0 7 8 2 | 1.  2 8 9 2 7 1 | 0  .  7 0 6 4 7 6 | 6  .  1 2 6 5 2 5 | 3  .  6 2 3 1 3 7 | 3.  2 2 2 4 3 0 | 3.  5 6 8 1 3 6 |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | m  i  n | 1  8  .  0  0 0 0  0 0 | 1 0  2  .  0 0 0 0 0 0 | 1  .  0 0 0 0 0 0 | 1  .  0 0 0 0 0 0 | 1  .  0 | 1  .  0 0 0 0 0 0 | 1.  0  0  0 0 0 0 | 3  0  .  0  0 0 0  0 0 | 1  .  0 0 0 0 0 0 | 1  .  0 0 0 0 0 0 | 1  .  0 0 0 0 0 0 | 1  0 0  9  .  0 0 0  0  0 0 | 2  0 9  4  .  0 0 0  0  0 0 | 0.  0  0  0 0 0 0 | 1  1  .  0  0 0 0  0 0 | 3  .  0 0 0 0 0 0 | 1.  0  0  0 0 0 0 | 8  0  .  0 | 0  .  0 0 0 0 0 0 | 0  .  0 0 0 0 0 0 | 0.  0  0  0 0 0 0 | 1  .  0 0 0 0 0 0 | 0  .  0 0 0 0 0 0 | 0  .  0 0 0 0 0 0 | 0.  0  0  0 0 0 0 | 0.  0  0  0 0 0 0 |
| 2 5  % | 3  0  .  0 0 0 0 0 0 | 4 6  5  .  0 0 0 0 0 0 | 2  .  0 0 0 0 0 0 | 2  .  0 0 0 0 0 0 | 1  .  0 | 4 9  1  .  2 5 0 0 0 0 | 2.  0 0 0 0 0 0 | 4  8  .  0 0 0 0 0 0 | 2  .  0 0 0 0 0 0 | 1  .  0 0 0 0 0 0 | 2  .  0 0 0 0 0 0 | 2 9 1  1  .  0 0 0 0 0 0 | 8 0 4  7  .  0 0 0 0 0 0 | 1.  0 0 0 0 0 0 | 1  2  .  0 0 0 0 0 0 | 3  .  0 0 0 0 0 0 | 2.  0 0 0 0 0 0 | 8  0  .  0 | 0  .  0 0 0 0 0 0 | 6  .  0 0 0 0 0 0 | 2.  0 0 0 0 0 0 | 2  .  0 0 0 0 0 0 | 3  .  0 0 0 0 0 0 | 2  .  0 0 0 0 0 0 | 0.  0 0 0 0 0 0 | 2.  0 0 0 0 0 0 |
| 5 | 3 | 8 | 7 | 3 | 1 | 1 | 3. | 6 | 3 | 2 | 3 | 4 | 1 | 2. | 1 | 3 | 3. | 8 | 1 | 1 | 3. | 3 | 5 | 3 | 1. | 3. |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | 0  % | 6  .  0 0  0  0 0 0 | 0  2  .  0  0  0 0 0 0 | .  0 0 0  0  0 0 | .  0 0 0  0  0 0 | .  0 | 0  2  0  .  5  0 0 0  0 0 | 0  0 0 0  0 0 | 6  .  0 0  0  0 0 0 | .  0 0 0  0  0 0 | .  0 0 0  0  0 0 | .  0 0 0  0  0 0 | 9  1  9  .  0  0 0 0  0 0 | 4  2 3  5  .  5 0 0  0  0 0 | 0  0 0 0  0 0 | 4  .  0 0  0  0 0 0 | .  0 0 0  0  0 0 | 0  0 0 0  0 0 | 0  .  0 | .  0 0 0  0  0 0 | 0  .  0 0  0  0 0 0 | 0  0 0 0  0 0 | .  0 0 0  0  0 0 | .  0 0 0  0  0 0 | .  0 0 0  0  0 0 | 0  0 0 0  0 0 | 0  0 0 0  0 0 |
| 7 5  % | 4  3  .  0 0 0 0 0 0 | 1 1 5  7  .  0 0 0 0 0 0 | 1  4  .  0 0 0 0 0 0 | 4  .  0 0 0 0 0 0 | 1  .  0 | 1 5 5  5  .  7 5 0 0 0 0 | 4.  0 0 0 0 0 0 | 8  3  .  7 5 0 0 0 0 | 3  .  0 0 0 0 0 0 | 3  .  0 0 0 0 0 0 | 4  .  0 0 0 0 0 0 | 8 3 7  9  .  0 0 0 0 0 0 | 2 0 4 6  1  .  5 0 0 0 0 0 | 4.  0 0 0 0 0 0 | 1  8  .  0 0 0 0 0 0 | 3  .  0 0 0 0 0 0 | 4.  0 0 0 0 0 0 | 8  0  .  0 | 1  .  0 0 0 0 0 0 | 1  5  .  0 0 0 0 0 0 | 3.  0 0 0 0 0 0 | 3  .  0 0 0 0 0 0 | 9  .  0 0 0 0 0 0 | 7  .  0 0 0 0 0 0 | 3.  0 0 0 0 0 0 | 7.  0 0 0 0 0 0 |
| m | 6 | 1 | 2 | 5 | 1 | 2 | 4. | 1 | 4 | 5 | 4 | 1 | 2 | 9. | 2 | 4 | 4. | 8 | 3 | 4 | 6. | 4 | 4 | 1 | 1 | 1 |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | |



|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  |  | i t i  o n | n e s s T r  a v e  l | y  R  a  t  e | r t  m  e n  t | a n c e F r  o  m H o  m  e | a t i  o n | a t i  o n F  i  e l  d | o y e e C o u n  t | o y e e N u  m  b e  r | o n  m  e n t  S a ti s  f  a c ti o n | e  r | l  y  R  a t  e | n v o  l  v e  m  e n  t | e v e  l | o l  e | a t i  s f  a c  t i  o n | t  a  l  S t  a  t  u s | h l  y  I  n c o  m  e | h l  y  R a t  e | o  m  p a n i  e s  W o  r  k e d | 1  8 | T i  m  e | e n  t  S a l  a  r  y  H  i  k e | o r  m  a n c e R a t i  n g | ti o n s h i  p S a ti s  f  a c ti o n | d a  r  d H o u  r  s | k  O  p  t i  o n L e v e  l | l  W  o  r  k i  n g Y e a  r  s | n i  n g T i  m  e s L a s t  Y e a  r | L i f  e B a  l  a n c e | s  A  t  C o  m  p a n y |  |
| 0 | 4  1 | Y e s | T  r  a v e  l  \_ R  a | 1 1 0 2 | S a  l  e s | 1 | 2 | L  i f  e S  c  i  e n | 1 | 1 | 2 | F e  m  a  l  e | 9  4 | 3 | 2 | S a  l  e s  E  x e c | 4 | S  i  n g  l  e | 5 9 9 3 | 1 9 4 7 9 | 8 | Y | Y e s | 1  1 | 3 | 1 | 8  0 | 0 | 8 | 0 | 1 | 6 |  |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  |  |  | r  e  l  y |  |  |  |  | c e s |  |  |  |  |  |  |  | 1. t i 2. e |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 4  9 | N  o | T  r  a v e  l  \_ F  r  e q u e n  t l  y | 2 7 9 | R  e s e a  r  c h  & D  e v e  l  o p  m  e n  t | 8 | 1 | L  i f  e S  c  i  e n c e s | 1 | 2 | 3 | M  a  l  e | 6  1 | 2 | 2 | R  e s e a  r  c h S  c  i  e n  t i  s  t | 2 | M  a  r r  i  e d | 5 1 3 0 | 2 4 9 0 7 | 1 | Y | N  o | 2  3 | 4 | 4 | 8  0 | 1 | 1  0 | 3 | 3 | 1  0 |  |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | 2 | 3  7 | Y e s | T r  a v e  l  \_ R a r  e  l  y | 1  3 7 3 | R e s e a  r  c h  & D  e v e l  o p  m  e n  t | 2 | 2 | O  t  h e  r | 1 | 4 | 4 | M a  l  e | 9  2 | 2 | 1 | L a b o r  a t  o r  y T e c h n  i  c  i  a n | 3 | S i  n g  l  e | 2  0 9 0 | 2  3 9 6 | 6 | Y | Y e s | 1  5 | 3 | 2 | 8  0 | 0 | 7 | 3 | 3 | 0 |  |
| 3 | 3  3 | N  o | T  r  a v e | 1 3 9 2 | R  e s e a | 3 | 4 | L  i f  e  S | 1 | 5 | 4 | F e  m  a  l | 5  6 | 3 | 1 | R  e s e a | 3 | M  a  r r  i | 2 9 0 9 | 2 3 1 5 | 1 | Y | Y e s | 1  1 | 3 | 3 | 8  0 | 0 | 8 | 3 | 3 | 8 |  |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  |  |  | l  \_ F  r  e q u e n t l  y |  | r  c h  & D  e v e l  o p  m e n  t |  |  | c i  e n c e s |  |  |  | e |  |  |  | r  c h S c  i  e n t i  s  t |  | e d |  | 9 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | 2  7 | N  o | T  r  a v e  l  \_ R  a  r  e | 5 9 1 | R  e s e a  r  c h  & D  e | 2 | 1 | M  e d  i  c a  l | 1 | 7 | 1 | M  a  l  e | 4  0 | 3 | 1 | L a b o  r  a  t  o  r  y  T | 2 | M  a  r r  i  e d | 3 4 6 8 | 1 6 6 3 2 | 9 | Y | N  o | 1  2 | 3 | 4 | 8  0 | 1 | 6 | 3 | 3 | 2 |  |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

l

y

v

e

l

o

p

m

e

n

t

e

c

h

n

i

c

i

a

n

Attrition

Rate

Attrition

rate:

The

attrition

rate

measures

the

percentage

of

employees

who

leave

the

company

in

a

given

period

of

time.

It

is

usually

calculated

within

a

year

and

is

expressed

as

a

percentage

of

the

total

number

of

employees.

In

[9]:

df['Attrition'].value\_counts(normalize=True)

Out[9]:

Attrition

No

0.838776

Yes

0.161224

Name:

proportion,

dtype:

float64

The

output

displays

the

proportion

of

employees

with

regard

to

attrition

status

in

the

|  |
| --- |
| dataset. Let's visualize it!  In [10]:  attrition = df['Attrition'].value\_counts(normalize=True)  plt.figure(figsize=(8,6)) ax = sns.barplot(x=attrition.index, y=attrition)  for p in ax.patches:  ax.annotate(f'{p.get\_height() \* 100:.2f}%',  (p.get\_x() + p.get\_width() / 2.,  p.get\_height()), ha='center', va='bottom')  plt.title('Distribution of Attrition Rate') plt.xlabel('Attrition') plt.ylabel('Percentage') plt.tight\_layout() plt.show() |

Based

on

the

analysis,

the

company's

attrition

rate

is

16.12

%.

This

means

that

about

16.12

%

of

the

employees

decided

to

leave

the

company

during

the

analyzed

period.

Average

of

Tenure

Average

tenure:

The

average

tenure

measures

the

average

number

of

years

an

employee

stays

with

the

company

before

leaving.

It

can

provide

insight

into

workforce

stability

and

employee

satisfaction

within

the

organization.

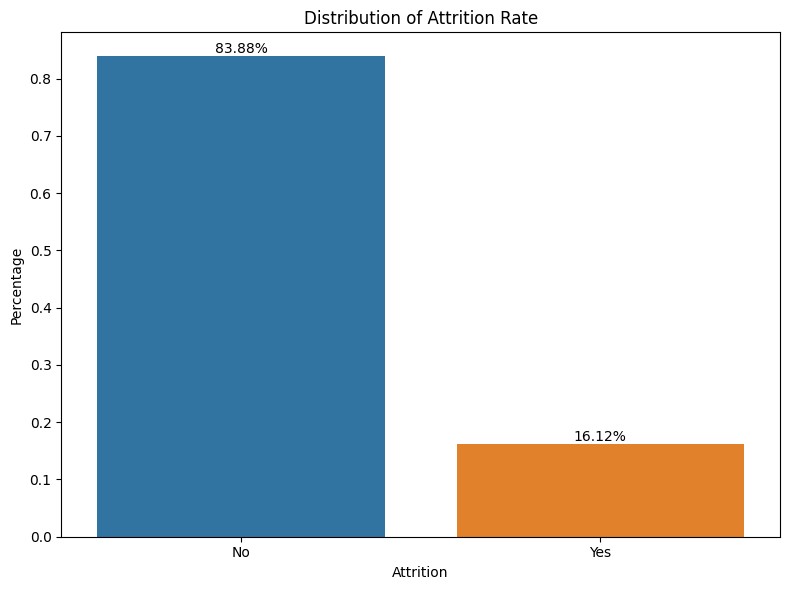
In

[11]:

avg\_tenure

=

df['YearsAtCompany'].mean()



|  |
| --- |
| print(f'Average years of employee to leave the company is {avg\_tenure} years')  Average years of employee to leave the company is  7.0081632653061225 years  The average tenure of employees before they decided to leave was 7.01 years. With this average tenure, it can be concluded that many employees feel comfortable and have been with the company for a long time.  Employee's Demographics  In [12]: fig, axes = plt.subplots(nrows=1, ncols=3, figsize=(15,5))  sns.histplot(data=df, x='Age', kde=True, ax=axes[0]) axes[0].set\_title('Distribution Employee by Age') axes[0].set\_xlabel('Age') axes[0].set\_ylabel('Count')  sns.countplot(data=df, x='Gender', ax=axes[1]) axes[1].set\_title('Distribution Employee by Gender') axes[1].set\_xlabel('Gender') axes[1].set\_ylabel('Count') |

sns.countplot(data=df,

x='Department',

ax=axes[2])

axes[2].set\_title('Distribution

Employee

by

Department')

axes[2].set\_xlabel('Department')

axes[2].set\_ylabel('Count')

plt.tight\_layout()

plt.show()

1.

Age:

Most

of

the

company's

employees

are

in

the

30-35

age

group.

This

indicates

that

the

company

has

many

employees

who

are

at

a

productive

and

experienced

age.

2.

Gender:

The

majority

of

employees

at

this

company

are

male.

There

are

significantly

more

male

employees

than

female

employees.

3.

Department:

Most

of

the

company's

employees

are

concentrated

in

the

research

and

development

department.

This

indicates

that

the

company

is

heavily

focused

on

product

or

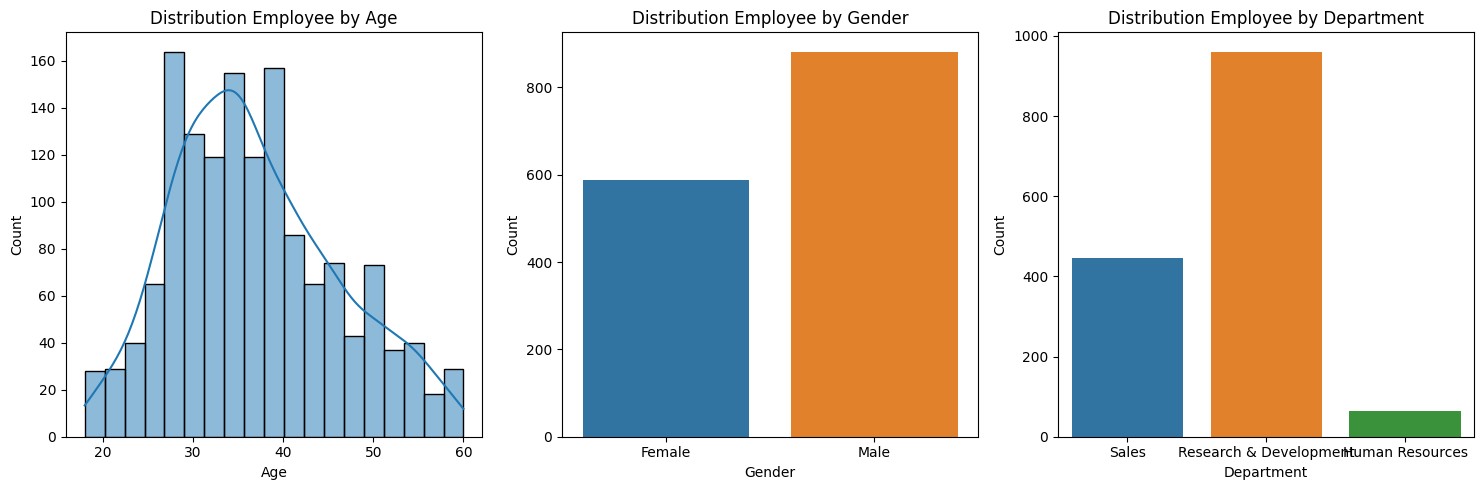
service

research

and

development

activities.



In [13]:

B

s

e

r

v

l

D

a

e

t

D

i

t

c

o

H

m

E

u

t

F

e

d

E

p

y

u

b

r

E

n

v

o

e

a

s

a

ti

e

r

H

r

a

e

J

b

v

e

e

t

b

e

N

u

m

o

a

s

o

k

d

v

8

e

m

R

e

l

ti

s

S

ti

f

c

S

t

c

p

n

e

e

T

r

a

n

g

e

L

s

Y

Y

a

A

m

a

y

Y

e

a

s

C

e

t

o

e

df\_attrition = df[df['Attrition'] == 'Yes'] df\_attrition.head() Out[13]:

P P T

J e e o W

E o M M r r S t o u s d m ir o C a o i e m b a o M c f t a r

A D E p S r n o e o a l k i e a c l n o I J m O o k i r

t a d l J a i t n n r n W L

n p n a o m G u n o p O v n O n s

t i u o o t t h t t m d o i

A r l c y b i a l h S a a r f s r e i e n n l o L n e r h t T t

g i y a e R s l y l a n r k e s t F o e t d y l e i r T i i i C

e t R t e o f S I y l c d i B

T m r n N S e R v v e 1 i p o m o i a i C l a t n R a e H n a

o t o o e c a c a r R o g l a n m i m ti t m l W e a L s p

n e n u t t o t y a u Y a

n i u m e H t r e n

e o l e f n r s v a n t o s e i i s a c

n k n r e e c e a l t

e g s

o ti e

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  |  |  |  |  |  |  |  |  |  |  | n |  |  |  |  |  |  |  |  |  |  |  |  |  |  | o n |  |  |  | a  r |  |  |  |
| 0 | 4  1 | Y e s | T r  a v e l  \_ R  a  r  e  l  y | 1  1  0 2 | S a  l  e s | 1 | 2 | L i f  e S c i  e n c e s | 1 | 1 | 2 | F e  m a  l  e | 9  4 | 3 | 2 | S a l  e s  E x e c u  t i  v e | 4 | S  i  n g  l  e | 5  9  9 3 | 1  9 4 7 9 | 8 | Y | Y e s | 1  1 | 3 | 1 | 8  0 | 0 | 8 | 0 | 1 | 6 |  |
| 2 | 3  7 | Y e s | T  r  a v e  l  \_  R | 1 3 7 3 | R  e s e a  r  c h | 2 | 2 | O  t  h e  r | 1 | 4 | 4 | M  a  l  e | 9  2 | 2 | 1 | L a b o  r  a  t  o | 3 | S  i  n g  l  e | 2 0 9 0 | 2 3 9 6 | 6 | Y | Y e s | 1  5 | 3 | 2 | 8  0 | 0 | 7 | 3 | 3 | 0 |  |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

4

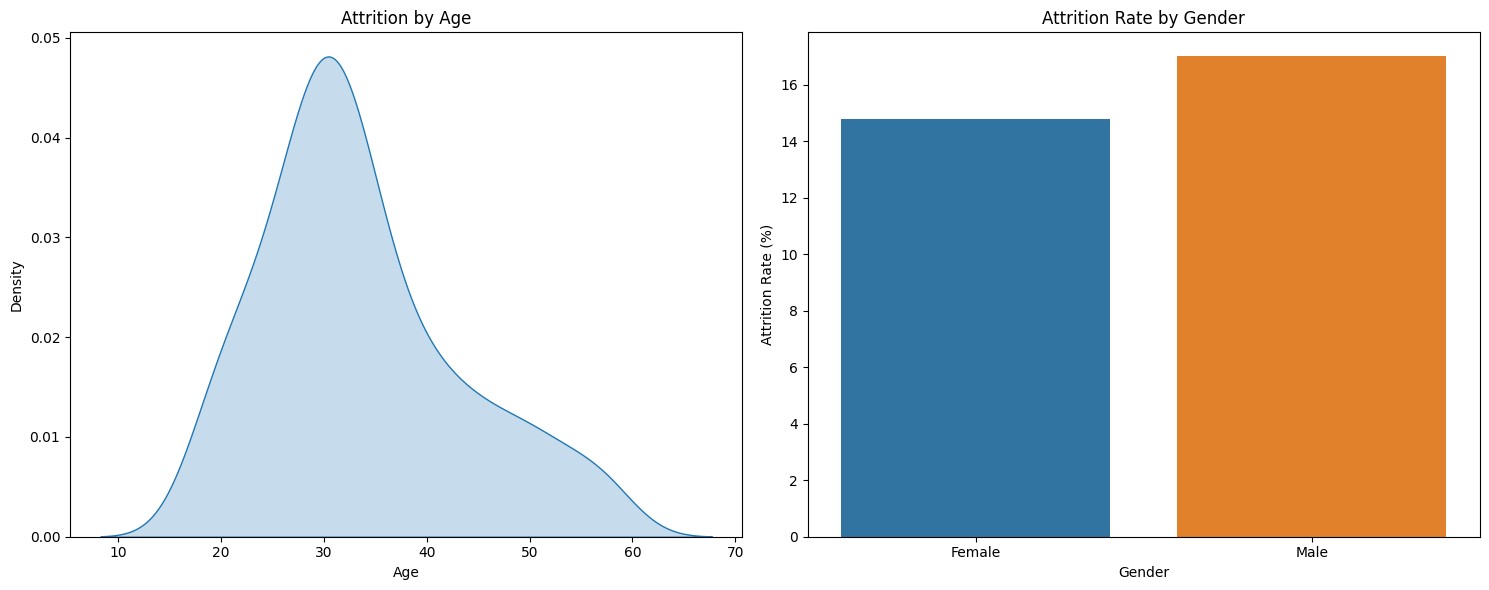
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  |  |  | a r  e  l  y |  | & D  e v e  l  o p  m e n  t |  |  |  |  |  |  |  |  |  |  | r  y T e c h n  i  c i  a n |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1  4 | 2  8 | Y e s | T  r  a v e  l  \_ R  a  r  e  l  y | 1 0 3 | R  e s e a  r  c h  & D  e v e  l | 2  4 | 3 | L  i f  e S  c  i  e n c e s | 1 | 1  9 | 3 | M  a  l  e | 5  0 | 2 | 1 | L a b o  r  a  t  o  r  y T e c h | 3 | S  i  n g  l  e | 2 0 2 8 | 1 2 9 4 7 | 5 | Y | Y e s | 1  4 | 3 | 2 | 8  0 | 0 | 6 | 4 | 3 | 4 |  |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  |  |  |  |  | o p  m  e n  t |  |  |  |  |  |  |  |  |  |  | n  i  c  i  a n |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2  1 | 3  6 | Y e s | T  r  a v e  l  \_ R  a  r  e  l  y | 1 2 1 8 | S a  l  e s | 9 | 4 | L  i f  e S  c  i  e n c e s | 1 | 2  7 | 3 | M  a  l  e | 8  2 | 2 | 1 | S a  l  e s  R  e p  r  e s e n  t  a  t i  v e | 1 | S  i  n g  l  e | 3 4 0 7 | 6 9 8 6 | 7 | Y | N  o | 2  3 | 4 | 2 | 8  0 | 0 | 1  0 | 4 | 3 | 5 |  |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | 2  4 | 3  4 | Y e s | T r  a v e  l  \_ R a r  e  l  y | 6 9 9 | R  e s e a r  c h  & D  e v e  l  o p  m  e n  t | 6 | 1 | M e d  i  c a  l | 1 | 3  1 | 2 | M  a  l  e | 8  3 | 3 | 1 | R  e s e a r  c h S  c i  e n  t i  s  t | 1 | S  i  n g  l  e | 2 9 6 0 | 1 7 1 0 2 | 2 | Y | N  o | 1  1 | 3 | 3 | 8  0 | 0 | 8 | 2 | 3 | 4 |  |
| Demografic Factors  In [14]:  *# Fungsi untuk Menghitung Attrition Rate* def calculate\_attrition\_rate(df, column): | | | | | | | | | | | | | | | | | | | | | | | | | | |  | | | |  | | |

0 2 3 2

|  |
| --- |
| attrition\_counts = df.groupby([column,  'Attrition']).size().unstack(fill\_value=0) attrition\_rate = attrition\_counts['Yes'] /  attrition\_counts.sum(axis=1) \* 100 attrition\_rate\_df = attrition\_rate.reset\_index() attrition\_rate\_df.columns = [column, 'AttritionRate'] return attrition\_rate\_df  In [15]: fig, axes = plt.subplots(nrows=1, ncols=2, figsize=(15,6))  *# Plot 1: KDE plot of Age with Attrition hue* sns.kdeplot(data=df\_attrition, x='Age', fill=True, ax=axes[0]) axes[0].set\_title('Attrition by Age') axes[0].set\_xlabel('Age') axes[0].set\_ylabel('Density')  *# Plot 2: Bar plot of Gender count with Attrition hue* attrition\_rate\_df = calculate\_attrition\_rate(df, 'Gender') sns.barplot(data=attrition\_rate\_df, x='Gender', y='AttritionRate', ax=axes[1]) axes[1].set\_title(f'Attrition Rate by Gender') axes[1].set\_xlabel('Gender') axes[1].set\_ylabel('Attrition Rate (%)') |

plt.tight\_layout()

plt.show()



[Reference](https://github.com/Niranjankumar-c/HRAnalyticsEmployeeAttrition) [link](https://github.com/Niranjankumar-c/HRAnalyticsEmployeeAttrition)