Internshala Assignment

IDZ Digital Private Limited

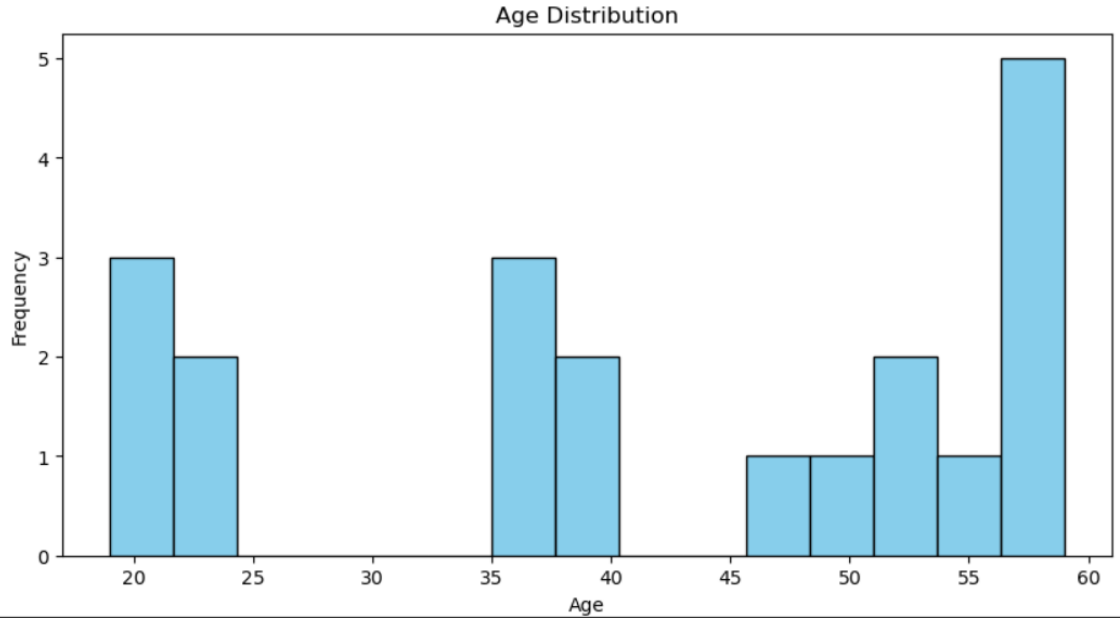
Data Analytics internship

Assignment: To find 20 interesting observations from this data and visualize the data.

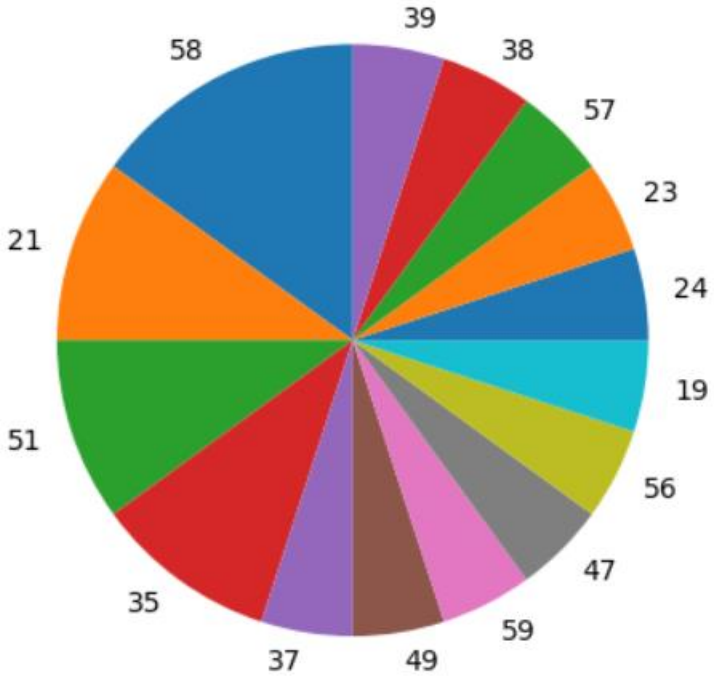
	SR.	first_name	Last_name	Sex	Age	Occupation	Salary	Marital Status	No. of Children
0	1	Lucy	Anderson	Female	58	Manager	4702	Married	3.0
1	2	Emily	Thomas	Female	21	Electrician	1943	Single	0.0
2	3	Edward	Stewart	Male	37	Producer	9026	Married	1.0
3	4	Frederick	Johnson	Male	49	Programmer	5998	Married	1.0
4	5	Lucy	Davis	Female	59	Pilot	8832	Married	3.0
5	6	Jenna	Ellis	Female	51	Baker	3578	Married	1.0
6	7	Adam	Nelson	Male	47	Photographer	2060	Married	1.0
7	8	Anna	Smith	Female	56	Fine Artist	9072	Married	3.0
8	9	Frederick	Kelly	Male	21	Chef	8115	Single	0.0
9	10	Tara	Brown	Female	19	Jeweller	4424	Single	0.0
10	11	Dominik	Martin	Male	58	Lecturer	4754	Married	3.0
11	12	Tara	Cameron	Female	24	Archeologist	6531	Married	3.0
12	13	Roman	Campbell	Male	23	Aeroplane Pilot	9654	Married	3.0
13	14	Alina	Russell	Female	51	Engineer	8411	Married	0.0
14	15	Alford	Myers	Male	57	Auditor	1181	Married	3.0
15	16	Preston	Mason	Male	38	Electrician	4817	Married	0.0
16	17	April	Stevens	Female	35	Singer	4837	Single	2.0
17	18	Aida	Gibson	Female	35	Archeologist	3403	Single	2.0
18	19	Dexter	Williams	Male	58	Insurer	7075	Married 0	NaN
19	20	Vivian	Taylor	Female	39	Veterinarian	1796	Married	3.0

Visualization of Age

```
plt.figure(figsize=(10, 5))
plt.hist(df['Age'], bins=15, color='skyblue', edgecolor='black')
plt.title('Age Distribution')
plt.xlabel('Age')
plt.ylabel('Frequency')
plt.show()
```

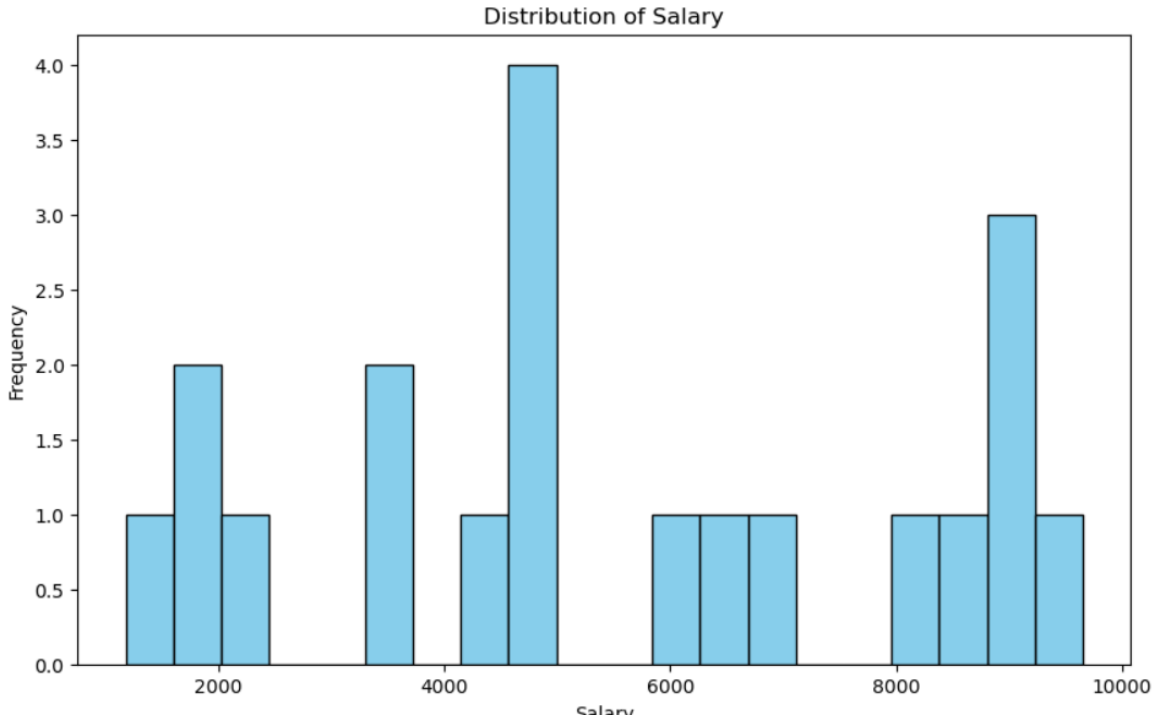



```
:  Age_counts = df['Age'].value_counts()
   plt.pie(Age_counts, labels=Age_counts.index, startangle=90)
   plt.show()
```

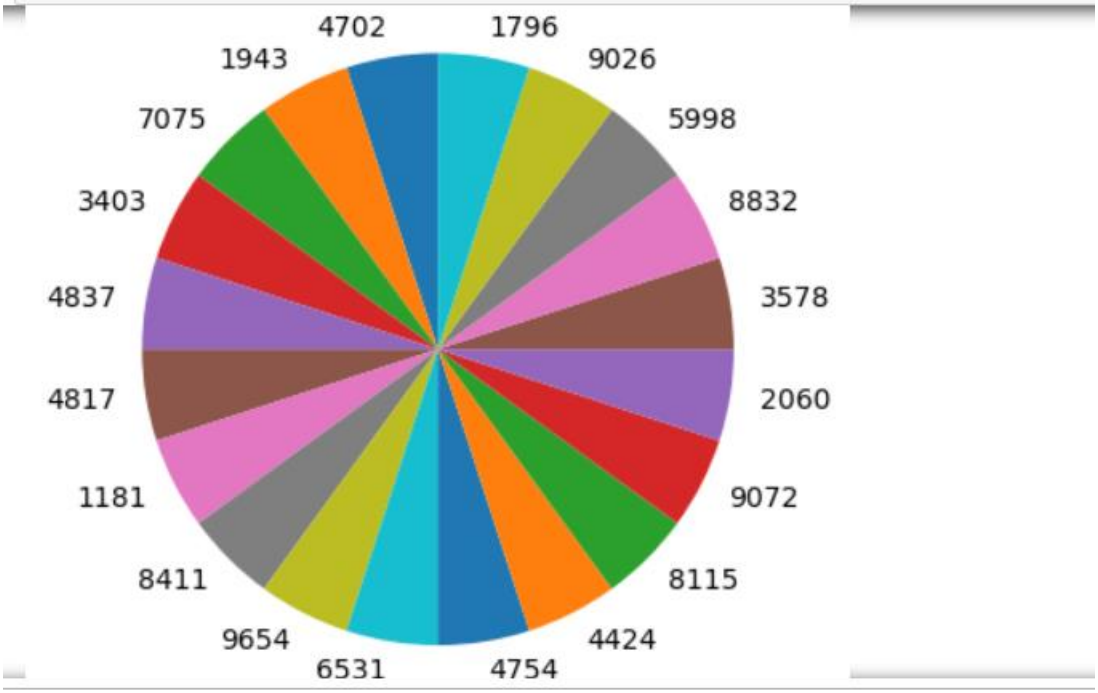


Visualization of Salary

```
plt.hist(df['Salary'], bins=20, color='skyblue', edgecolor='black')
plt.title('Distribution of Salary')
plt.xlabel('Salary')
plt.ylabel('Frequency')
plt.show()
```

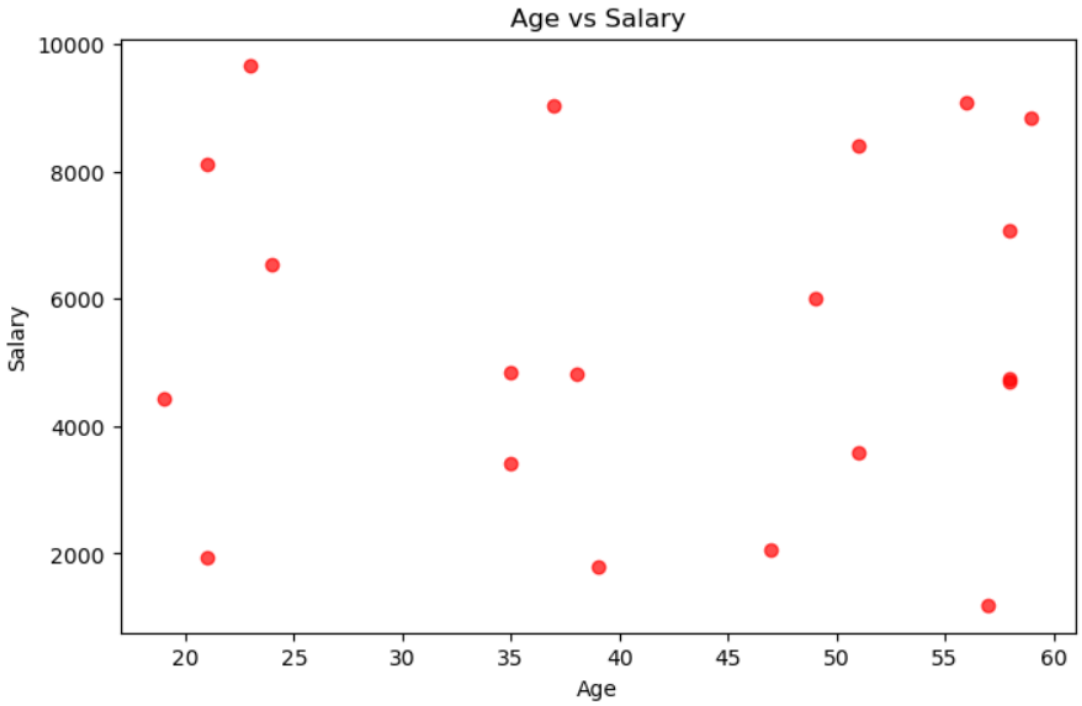



```
Salary = df['Salary'].value_counts()
plt.pie(Salary, labels=Salary.index, startangle=90)
plt.show()
```



Scatter Plot Between Salary and Age

```
plt.figure(figsize=(8, 5))
plt.scatter(df['Age'], df['Salary'], color='red', alpha=0.7)
plt.title('Age vs Salary')
plt.xlabel('Age')
plt.ylabel('Salary')
plt.show()
```



Scatter Plot Between No. of Childrens and Age

```
plt.figure(figsize=(10, 5))
plt.scatter(df['Age'], df['No. of Children'], color='green', alpha=0.7)
plt.title('Age vs Number of Children')
plt.xlabel('Age')
plt.ylabel('Number of Children')
plt.show()
```