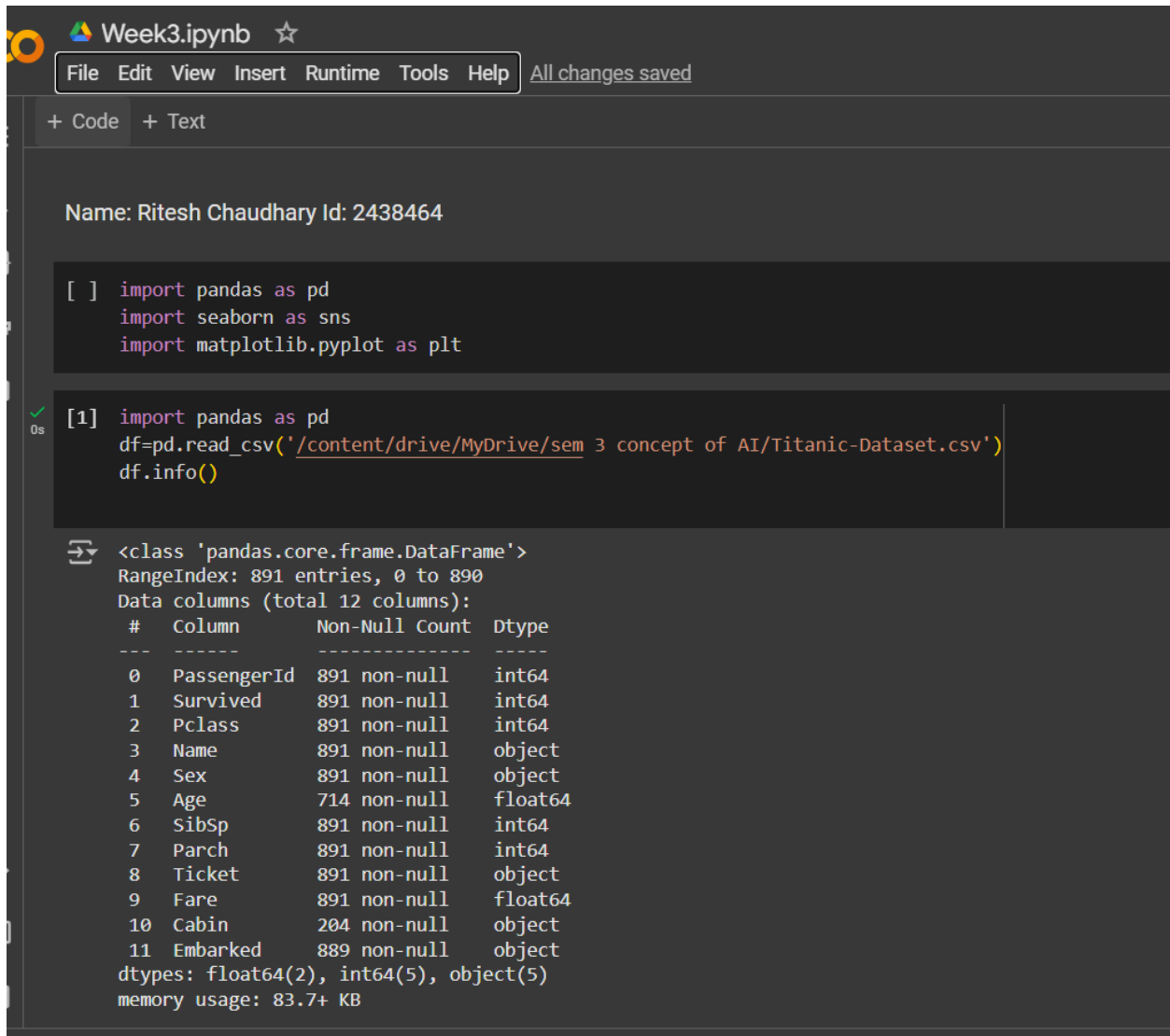


Name: Ritesh Chaudhary

Id: 2438464

## Workshop Week-3



The screenshot shows a Jupyter Notebook titled "Week3.ipynb" with a menu bar (File, Edit, View, Insert, Runtime, Tools, Help) and a status bar ("All changes saved"). The notebook contains two code cells. The first cell imports pandas, seaborn, and matplotlib. The second cell reads a CSV file from Google Drive and displays its information. The output of the second cell shows the DataFrame's structure, including column names, non-null counts, and data types.

Name: Ritesh Chaudhary Id: 2438464

```
[ ] import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

```
[1] import pandas as pd
df=pd.read_csv('/content/drive/MyDrive/sem 3 concept of AI/Titanic-Dataset.csv')
df.info()
```

<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 891 entries, 0 to 890  
Data columns (total 12 columns):

#	Column	Non-Null Count	Dtype
0	PassengerId	891 non-null	int64
1	Survived	891 non-null	int64
2	Pclass	891 non-null	int64
3	Name	891 non-null	object
4	Sex	891 non-null	object
5	Age	714 non-null	float64
6	SibSp	891 non-null	int64
7	Parch	891 non-null	int64
8	Ticket	891 non-null	object
9	Fare	891 non-null	float64
10	Cabin	204 non-null	object
11	Embarked	889 non-null	object

dtypes: float64(2), int64(5), object(5)  
memory usage: 83.7+ KB

memory usage: 83.7+ KB

```
[2] fare=df['Fare']  
print(fare.head())
```

```
0    7.2500  
1   71.2833  
2    7.9250  
3   53.1000  
4    8.0500  
Name: Fare, dtype: float64
```

```
class_age=df[['Pclass','Age']]  
print(class_age.head())
```

```
   Pclass  Age  
0        3  22.0  
1        1  38.0  
2        3  26.0  
3        1  35.0  
4        3  35.0
```

```
survived_gender=df[['Survived','Sex']]  
print(survived_gender.head())
```

```
   Survived  Sex  
0          0  male  
1          1 female  
2          1 female  
3          1 female  
4          0  male
```

+ Code + Text

✓  
0s



```
passanger_Fare_gt=df[df['Fare']>100]  
print(passanger_Fare_gt)
```



779	Robert, Mrs. Edward Scott (Elisabeth Walton Mc...	female	43.00	0
802	Carter, Master. William Thornton II	male	11.00	1
856	Wick, Mrs. George Dennick (Mary Hitchcock)	female	45.00	1

	Parch	Ticket	Fare	Cabin	Embarked
27	2	19950	263.0000	C23 C25 C27	S
31	0	PC 17569	146.5208	B78	C
88	2	19950	263.0000	C23 C25 C27	S
118	1	PC 17558	247.5208	B58 B60	C
195	0	PC 17569	146.5208	B80	C
215	0	35273	113.2750	D36	C
258	0	PC 17755	512.3292	NaN	C
268	1	PC 17582	153.4625	C125	S
269	0	PC 17760	135.6333	C99	S
297	2	113781	151.5500	C22 C26	S
299	1	PC 17558	247.5208	B58 B60	C
305	2	113781	151.5500	C22 C26	S
306	0	17421	110.8833	NaN	C
307	0	PC 17758	108.9000	C65	C
311	2	PC 17608	262.3750	B57 B59 B63 B66	C
318	2	36928	164.8667	C7	S
319	1	16966	134.5000	E34	C
325	0	PC 17760	135.6333	C32	C
332	1	PC 17582	153.4625	C91	S
334	0	PC 17611	133.6500	NaN	S
337	0	16966	134.5000	E40	C
341	2	19950	263.0000	C23 C25 C27	S
373	0	PC 17760	135.6333	NaN	C
377	2	113503	211.5000	C82	C
380	0	PC 17757	227.5250	NaN	C
390	2	113760	130.0000	B06 B08	S

+ Code + Text

```
Os 609 0 PC 17582 153.4625 C125 S
659 2 35273 113.2750 D48 C
660 0 PC 17611 133.6500 NaN S
679 1 PC 17755 512.3292 B51 B53 B55 C
689 1 24160 211.3375 B5 S
698 1 17421 110.8833 C68 C
700 0 PC 17757 227.5250 C62 C64 C
708 0 113781 151.5500 NaN S
716 0 PC 17757 227.5250 C45 C
730 0 24160 211.3375 B5 S
737 0 PC 17755 512.3292 B101 C
742 2 PC 17608 262.3750 B57 B59 B63 B66 C
763 2 113760 120.0000 B96 B98 S
779 1 24160 211.3375 B3 S
802 2 113760 120.0000 B96 B98 S
856 1 36928 164.8667 NaN S
```

```
Os 6 filtered_Pclass=df[df['Pclass']==1]
print(filtered_Pclass)
```

```
PassengerId Survived Pclass \
1 2 1 1
3 4 1 1
6 7 0 1
11 12 1 1
23 24 1 1
.. ... ...
871 872 1 1
872 873 0 1
879 880 1 1
887 888 1 1
889 890 1 1
```

Name Sex Age SibSp \

```
Name Sex Age SibSp \
1 Cumings, Mrs. John Bradley (Florence Briggs Th... female 38.0 1
3 Futrelle, Mrs. Jacques Heath (Lily May Peel) female 35.0 1
6 McCarthy, Mr. Timothy J male 54.0 0
11 Bonnell, Miss. Elizabeth female 58.0 0
23 Sloper, Mr. William Thompson male 28.0 0
.. ... ...
871 Beckwith, Mrs. Richard Leonard (Sallie Monypeny) female 47.0 1
872 Carlsson, Mr. Frans Olof male 33.0 0
879 Potter, Mrs. Thomas Jr (Lily Alexenia Wilson) female 56.0 0
887 Graham, Miss. Margaret Edith female 19.0 0
889 Behr, Mr. Karl Howell male 26.0 0
```

```
Parch Ticket Fare Cabin Embarked
1 0 PC 17599 71.2833 C85 C
3 0 113803 53.1000 C123 S
6 0 17463 51.8625 E46 S
11 0 113783 26.5500 C103 S
23 0 113788 35.5000 A6 S
.. ... ...
871 1 11751 52.5542 D35 S
872 0 695 5.0000 B51 B53 B55 S
879 1 11767 83.1583 C50 C
887 0 112053 30.0000 B42 S
889 0 111369 30.0000 C148 C
```

[216 rows x 12 columns]

+ Code + Text

✓  
0s

```
[7] passangerAge=df[(df['Age']<18) & (df['Sex']=='female')]  
print(passangerAge.head())
```

⇒

	PassengerId	Survived	Pclass	Name \
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)
10	11	1	3	Sandstrom, Miss. Marguerite Rut
14	15	0	3	Vestrom, Miss. Hulda Amanda Adolfina
22	23	1	3	McGowan, Miss. Anna "Annie"
24	25	0	3	Palsson, Miss. Torborg Danira

	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
9	female	14.0	1	0	237736	30.0708	NaN	C
10	female	4.0	1	1	PP 9549	16.7000	G6	S
14	female	14.0	0	0	350406	7.8542	NaN	S
22	female	15.0	0	0	330923	8.0292	NaN	Q
24	female	8.0	3	1	349909	21.0750	NaN	S

✓  
0s

```
embarked_c_or_s=df[(df['Embarked']=='C') | (df['Embarked']=='S')]  
print(embarked_c_or_s.head())
```

⇒

	PassengerId	Survived	Pclass	\
0	1	0	3	
1	2	1	1	
2	3	1	3	
3	4	1	1	
4	5	0	3	

	Name	Sex	Age	SibSp	\
0	Braund, Mr. Owen Harris	male	22.0	1	
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	
2	Heikkinen, Miss. Laina	female	26.0	0	
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	
4	Allen, Mr. William Henry	male	25.0	0	

```
filtered_Pclass=df[(df['Pclass'].isin([1,2]))]
print(filtered_Pclass)
```

	PassengerId	Survived	Pclass	\
1	2	1	1	
3	4	1	1	
6	7	0	1	
9	10	1	2	
11	12	1	1	
..	...	...	...	
880	881	1	2	
883	884	0	2	
886	887	0	2	
887	888	1	1	
889	890	1	1	

	Name	Sex	Age	SibSp	\
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	
6	McCarthy, Mr. Timothy J	male	54.0	0	
9	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	
11	Bonnell, Miss. Elizabeth	female	58.0	0	
..	...	...	...	...	
880	Shelley, Mrs. William (Imanita Parrish Hall)	female	25.0	0	
883	Banfield, Mr. Frederick James	male	28.0	0	
886	Montvila, Rev. Juozas	male	27.0	0	
887	Graham, Miss. Margaret Edith	female	19.0	0	
889	Behr, Mr. Karl Howell	male	26.0	0	

	Parch	Ticket	Fare	Cabin	Embarked
1	0	PC 17599	71.2833	C85	C
3	0	113803	53.1000	C123	S
6	0	17463	51.8625	E46	S

+ Code + Text

```
[10] df['Fare_per_year']=df['Fare']/df['Age']  
print(df.head())
```

	PassengerId	Survived	Pclass	\
0	1	0	3	
1	2	1	1	
2	3	1	3	
3	4	1	1	
4	5	0	3	

	Name	Sex	Age	SibSp	\
0	Braund, Mr. Owen Harris	male	22.0	1	
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	
2	Heikkinen, Miss. Laina	female	26.0	0	
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	
4	Allen, Mr. William Henry	male	35.0	0	

	Parch	Ticket	Fare	Cabin	Embarked	Fare_per_year
0	0	A/5 21171	7.2500	NaN	S	0.329545
1	0	PC 17599	71.2833	C85	C	1.875876
2	0	STON/O2. 3101282	7.9250	NaN	S	0.304808
3	0	113803	53.1000	C123	S	1.517143
4	0	373450	8.0500	NaN	S	0.230000

```
[11] high=df[df['Fare_per_year']>5]  
print(high)
```

	PassengerId	Survived	Pclass	Name	\
7	8	0	3	Palsson, Master. Gosta Leonard	
16	17	0	3	Rice, Master. Eugene	
27	28	0	1	Fortune, Mr. Charles Alexander	
43	44	1	2	Laroche, Miss. Simonne Marie Anne Andree	

```
high=df[df['Fare_per_year']>5]
print(high)
```

	PassengerId	Survived	Pclass	Name	\
7	8	0	3	Palsson, Master. Gosta Leonard	
16	17	0	3	Rice, Master. Eugene	
27	28	0	1	Fortune, Mr. Charles Alexander	
43	44	1	2	Laroche, Miss. Simonne Marie Anne Andree	
50	51	0	3	Panula, Master. Juha Niilo	
..	...	...	...	...	
813	814	0	3	Andersson, Miss. Ebba Iris Alfrida	
824	825	0	3	Panula, Master. Urho Abraham	
827	828	1	2	Mallet, Master. Andre	
831	832	1	2	Richards, Master. George Sibley	
850	851	0	3	Andersson, Master. Sigvard Harald Elias	

	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	\
7	male	2.00	3	1	349909	21.0750	NaN	
16	male	2.00	4	1	382652	29.1250	NaN	
27	male	19.00	3	2	19950	263.0000	C23 C25 C27	
43	female	3.00	1	2	SC/Paris 2123	41.5792	NaN	
50	male	7.00	4	1	3101295	39.6875	NaN	
..	...	...	...	...	...	...	...	
813	female	6.00	4	2	347082	31.2750	NaN	
824	male	2.00	4	1	3101295	39.6875	NaN	
827	male	1.00	0	2	S.C./PARIS 2079	37.0042	NaN	
831	male	0.83	1	1	29106	18.7500	NaN	
850	male	4.00	4	2	347082	31.2750	NaN	

	Embarked	Fare_per_year
7	S	10.537500
16	Q	14.562500
27	S	13.842105
43	C	13.859733



+ Code + Text

```
✓ [12] high_fare_age_srt=df.sort_values(by='Fare_per_year',ascending=False)
0s print(high_fare_age_srt.head())
```

```
➡
```

	PassengerId	Survived	Pclass	Name	Sex	\
	305	306	1	1	Allison, Master. Hudson Trevor	male
	297	298	0	1	Allison, Miss. Helen Loraine	female
	386	387	0	3	Goodwin, Master. Sidney Leonard	male
	164	165	0	3	Panula, Master. Eino Viljami	male
	183	184	1	2	Becker, Master. Richard F	male

	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	Fare_per_year
305	0.92	1	2	113781	151.5500	C22 C26	S	164.728261
297	2.00	1	2	113781	151.5500	C22 C26	S	75.775000
386	1.00	5	2	CA 2144	46.9000	NaN	S	46.900000
164	1.00	4	1	3101295	39.6875	NaN	S	39.687500
183	1.00	2	1	230136	39.0000	F4	S	39.000000

```
✓ [13] result=high_fare_age_srt[['Fare_per_year','Name']]
0s print(result)
```

```
➡
```

	Fare_per_year	Name
305	164.728261	Allison, Master. Hudson Trevor
297	75.775000	Allison, Miss. Helen Loraine
386	46.900000	Goodwin, Master. Sidney Leonard
164	39.687500	Panula, Master. Eino Viljami
183	39.000000	Becker, Master. Richard F
..	...	...
859	NaN	Razi, Mr. Raihed
863	NaN	Sage, Miss. Dorothy Edith "Dolly"
868	NaN	van Melkebeke, Mr. Philemon
878	NaN	Laleff, Mr. Kristo
888	NaN	Johnston, Miss. Catherine Helen "Carrie"

[801 rows x 2 columns]

✓  
0s [14] total=df['Fare'].sum()  
print("Total Fare:",total)

⇒ Total Fare: 28693.9493

✓  
0s [15] firstClass=df[df['Pclass']==1]['Fare']  
totalf=firstClass.sum()  
print(totalf)

⇒ 18177.4125

✓  
0s [18] secondClass=df[df['Pclass']==2]['Fare']  
totals=secondClass.sum()  
print("Total of second class:",totals)  
thirdClass=df[df['Pclass']==3]['Fare']  
totalT=thirdClass.sum()  
print("Total of third class:",totalT)

⇒ Total of second class: 3801.8417  
Total of third class: 6714.6951

✓  
0s [19] lst=[total,totals,totalT]  
print(lst)

⇒ [28693.9493, 3801.8417, 6714.6951]

```
[20] proptionfirst=totalf/total
proptionsecond=totals/total
proptionthird=totalT/total
print("First:",proptionfirst)
print("Second:",proptionsecond)
print("Third:",proptionthird)
```

First: 0.6334928771899656  
Second: 0.1324962855496507  
Third: 0.23401083726038366

```
[21] df['ageGroup']=df['Age'].apply(lambda Age:'child' if Age<18 else 'adult' if 18<= Age<=65 else 'senior' if Age>65 else 'invalid')
print(df.head())
```

	PassengerId	Survived	Pclass	\
0	1	0	3	
1	2	1	1	
2	3	1	3	
3	4	1	1	
4	5	0	3	

	Name	Sex	Age	SibSp	\
0	Braund, Mr. Owen Harris	male	22.0	1	
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	
2	Heikkinen, Miss. Laina	female	26.0	0	
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	
4	Allen, Mr. William Henry	male	35.0	0	

	Parch	Ticket	Fare	Cabin	Embarked	Fare_per_year	ageGroup
0	0	A/5 21171	7.2500	NaN	S	0.329545	adult
1	0	PC 17599	71.2833	C85	C	1.875876	adult
2	0	STON/O2. 3101282	7.9250	NaN	S	0.304808	adult

+ Code + Text

✓  
0s [22] total\_passenger=len(df)  
print(total\_passenger)

↔ 891

✓  
0s [23] numberChild=len(df[df['ageGroup']=='child'])  
numberAdult=len(df[df['ageGroup']=='adult'])  
numberSenior=len(df[df['ageGroup']=='senior'])  
print("Number of children:",numberChild)  
print("Number of adult:",numberAdult)  
print("Number of senior:",numberSenior)

↔ Number of children: 113  
Number of adult: 593  
Number of senior: 8

✓  
0s [24] countChildprop= numberChild/total\_passenger  
countAdultprop=numberAdult/total\_passenger  
countSeniorprop=numberSenior/total\_passenger  
print("Proportion of children:",countChildprop)  
print("Proportion of adult:",countAdultprop)  
print("Proportion of senior:",countSeniorprop)


↔ Proportion of children: 0.12682379349046016  
Proportion of adult: 0.6655443322109988  
Proportion of senior: 0.008978675645342313

+ Code + Text

✓  
0s [23] Number of children: 113  
Number of adult: 593  
Number of senior: 8

✓  
0s [24] countChildprop= numberChild/total\_passenger  
countAdultprop=numberAdult/total\_passenger  
countSeniorprop=numberSenior/total\_passenger  
print("Proportion of children:",countChildprop)  
print("Proportion of adult:",countAdultprop)  
print("Proportion of senior:",countSeniorprop)

Proportion of children: 0.12682379349046016  
Proportion of adult: 0.6655443322109988  
Proportion of senior: 0.008978675645342313

✓  
0s  print("Proportion of children:",countChildprop,"%")  
print("Proportion of adult:",countAdultprop,"%")  
print("Proportion of senior:",countSeniorprop,"%")

Proportion of children: 0.12682379349046016 %  
Proportion of adult: 0.6655443322109988 %  
Proportion of senior: 0.008978675645342313 %