**Лабораторная работа № 1**

import matplotlib.pyplot as plt

import pandas as pd

df = pd.read\_csv('train.csv', ',')

pvt = df.pivot\_table('PassengerId', 'Pclass', 'Survived', 'count')

pvt.plot(kind = 'bar')

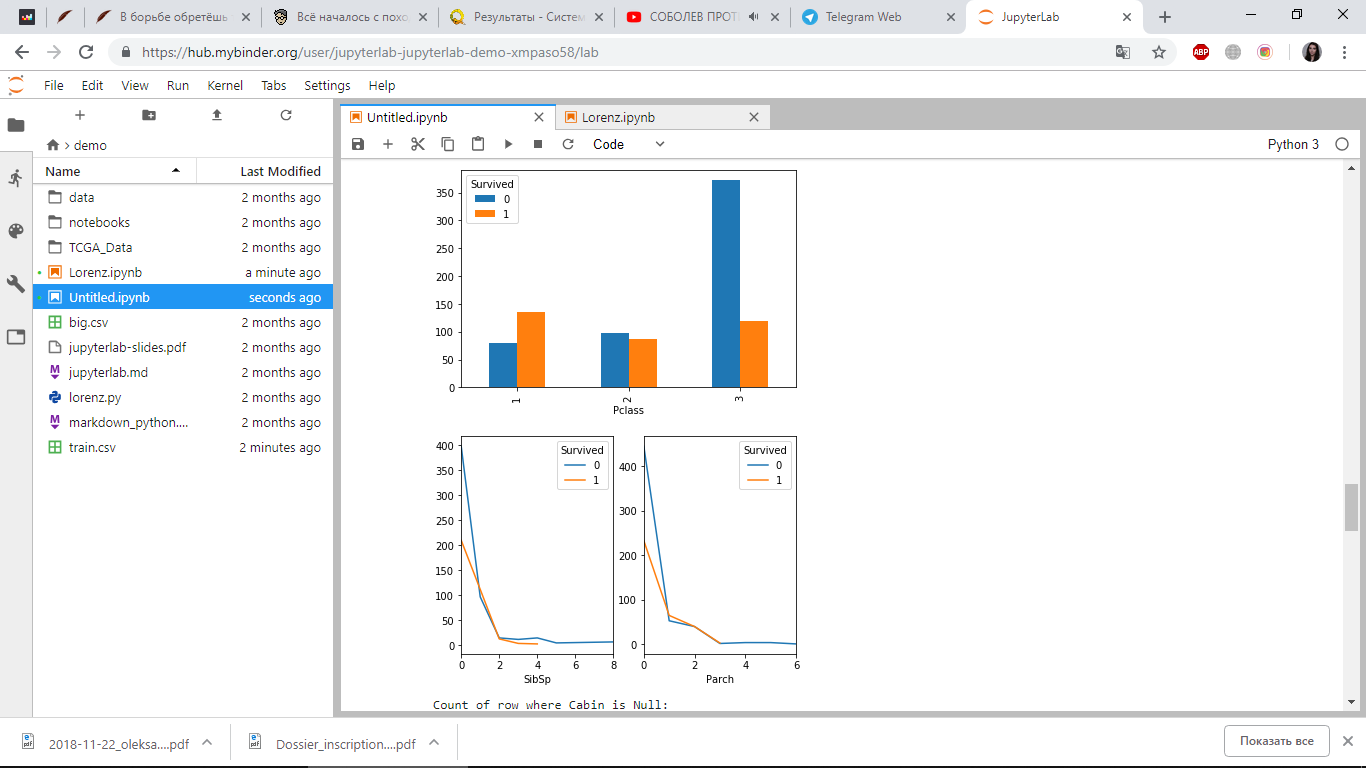
plt.show()

fig, axes = plt.subplots(ncols = 2)

df.pivot\_table('PassengerId', 'SibSp', 'Survived', 'count').plot(ax = axes[0])

df.pivot\_table('PassengerId', 'Parch', 'Survived', 'count').plot(ax = axes[1])

plt.show()



print('Count of row where Cabin is Null:')

print(df.PassengerId[df.Cabin.notnull()].count())

print('Count of row where Age is Null:')

print(df.PassengerId[df.Age.notnull()].count())

df.Age[df.Age.isnull()] = df.Age.median()

MaxPassEmbarked = df.groupby('Embarked').count()['PassengerId']

df.Embarked[df.Embarked.isnull()] = MaxPassEmbarked[MaxPassEmbarked == MaxPassEmbarked.max()].index[0]

print('Count of row where Fare is Null:')

print(df.PassengerId[df.Fare.isnull()].count())

df = df.drop(['PassengerId','Name','Ticket','Cabin'], axis=1)

Count of row where Cabin is Null:

204

Count of row where Age is Null:

714

Count of row where Fare is Null:

0

#Вариант 1

print('Количество пассажиров мужского пола из 1 класса:')

print(df.Pclass[df.Sex == 'male'][df.Pclass == 1].count())

Количество пассажиров мужского пола из 1 класса:

122

#Вариант 2

print('Количество детей (младше 14) из 2 класса:')

print(df.Pclass[df.Age <= 14][df.Pclass == 1].count())

Количество детей (младше 14) из 2 класса:

5

#Вариант 3

print('Количество одиноких пассажиров:')

print(df.Pclass[df.Parch == 0].count())

Количество одиноких пассажиров:

678

#Вариант 4

print('количество пассажиров, севших в порту Queenstown с дорогими билетами (ценна выше средней):')

print(df.Fare[df.Embarked == 'Q'][df.Fare > df.Fare.median()].count())

количество пассажиров, севших в порту Queenstown с дорогими билетами (ценна выше средней):

20

#Вариант 5

print('Cредний возраст пассажиров женского пола:')

print(df.Age[df.Sex == 'female'].median())

Cредний возраст пассажиров женского пола:

28.0

#Вариант 6

print('количество одиноких пожилых людей(старше 65):')

print(df.Age[df.Parch == 0][df.Age >= 65].count())

количество одиноких пожилых людей(старше 65):

9

#Вариант 7

print('Статистика по детям (младше 14 лет):')

print(df[df.Age <= 14])

Статистика по детям (младше 14 лет):

Survived Pclass Sex Age SibSp Parch Fare Embarked

7 0 3 male 2.00 3 1 21.0750 S

9 1 2 female 14.00 1 0 30.0708 C

10 1 3 female 4.00 1 1 16.7000 S

14 0 3 female 14.00 0 0 7.8542 S

16 0 3 male 2.00 4 1 29.1250 Q

24 0 3 female 8.00 3 1 21.0750 S

39 1 3 female 14.00 1 0 11.2417 C

43 1 2 female 3.00 1 2 41.5792 C

50 0 3 male 7.00 4 1 39.6875 S

58 1 2 female 5.00 1 2 27.7500 S

59 0 3 male 11.00 5 2 46.9000 S

63 0 3 male 4.00 3 2 27.9000 S

78 1 2 male 0.83 0 2 29.0000 S

119 0 3 female 2.00 4 2 31.2750 S

125 1 3 male 12.00 1 0 11.2417 C

147 0 3 female 9.00 2 2 34.3750 S

164 0 3 male 1.00 4 1 39.6875 S

165 1 3 male 9.00 0 2 20.5250 S

171 0 3 male 4.00 4 1 29.1250 Q

172 1 3 female 1.00 1 1 11.1333 S

182 0 3 male 9.00 4 2 31.3875 S

183 1 2 male 1.00 2 1 39.0000 S

184 1 3 female 4.00 0 2 22.0250 S

193 1 2 male 3.00 1 1 26.0000 S

205 0 3 female 2.00 0 1 10.4625 S

233 1 3 female 5.00 4 2 31.3875 S

237 1 2 female 8.00 0 2 26.2500 S

261 1 3 male 3.00 4 2 31.3875 S

278 0 3 male 7.00 4 1 29.1250 Q

297 0 1 female 2.00 1 2 151.5500 S

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

535 1 2 female 7.00 0 2 26.2500 S

541 0 3 female 9.00 4 2 31.2750 S

542 0 3 female 11.00 4 2 31.2750 S

549 1 2 male 8.00 1 1 36.7500 S

618 1 2 female 4.00 2 1 39.0000 S

634 0 3 female 9.00 3 2 27.9000 S

642 0 3 female 2.00 3 2 27.9000 S

644 1 3 female 0.75 2 1 19.2583 C

683 0 3 male 14.00 5 2 46.9000 S

686 0 3 male 14.00 4 1 39.6875 S

691 1 3 female 4.00 0 1 13.4167 C

720 1 2 female 6.00 0 1 33.0000 S

731 0 3 male 11.00 0 0 18.7875 C

750 1 2 female 4.00 1 1 23.0000 S

751 1 3 male 6.00 0 1 12.4750 S

755 1 2 male 0.67 1 1 14.5000 S

777 1 3 female 5.00 0 0 12.4750 S

780 1 3 female 13.00 0 0 7.2292 C

787 0 3 male 8.00 4 1 29.1250 Q

788 1 3 male 1.00 1 2 20.5750 S

802 1 1 male 11.00 1 2 120.0000 S

803 1 3 male 0.42 0 1 8.5167 C

813 0 3 female 6.00 4 2 31.2750 S

819 0 3 male 10.00 3 2 27.9000 S

824 0 3 male 2.00 4 1 39.6875 S

827 1 2 male 1.00 0 2 37.0042 C

831 1 2 male 0.83 1 1 18.7500 S

850 0 3 male 4.00 4 2 31.2750 S

852 0 3 female 9.00 1 1 15.2458 C

869 1 3 male 4.00 1 1 11.1333 S

#Вариант 8

print('Средняя цена билета по портам:')

print(df.groupby('Embarked')['Fare'].median())

Средняя цена билета по портам:

Embarked

C 29.70

Q 7.75

S 13.00

#Вариант 9

print('Средняя цена билета для каждого социального класса:')

df.groupby('Pclass')['Fare'].median()

Средняя цена билета для каждого социального класса:

Pclass

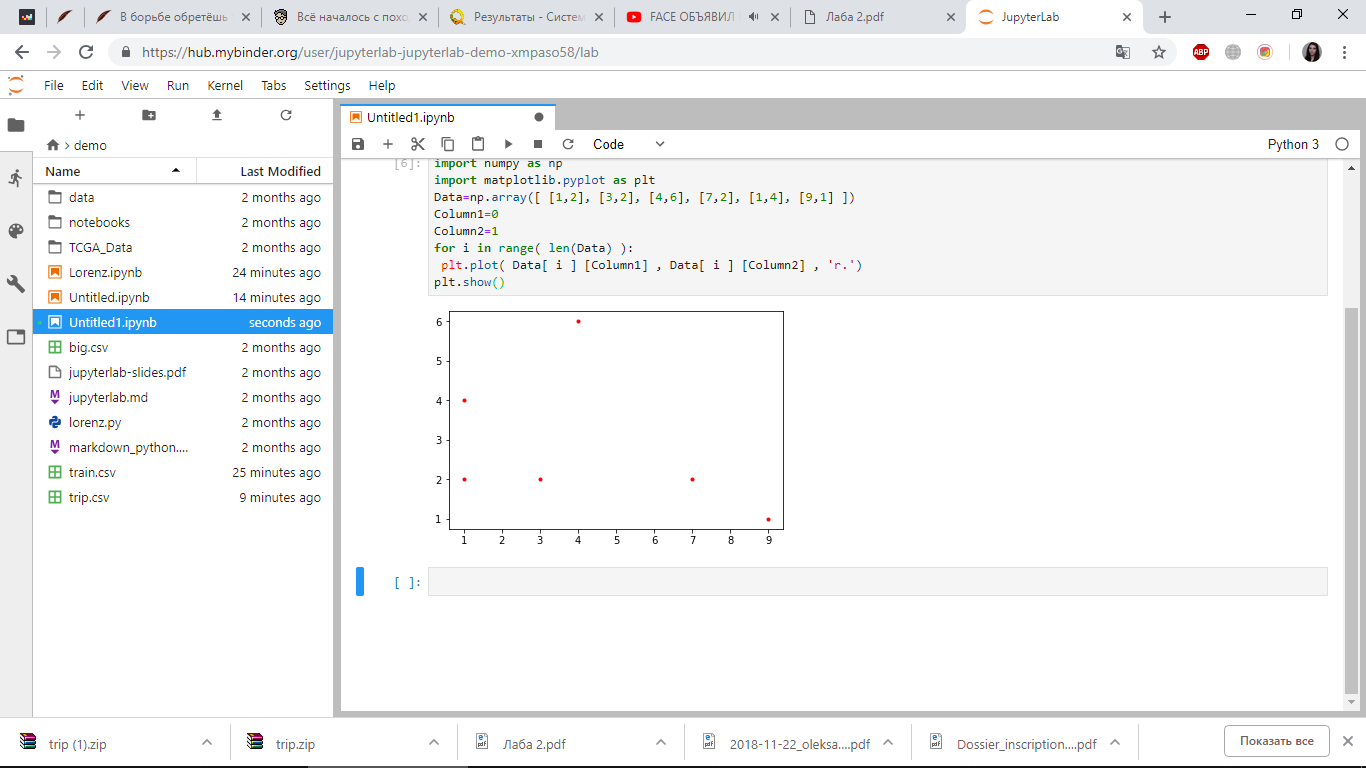
1 60.2875

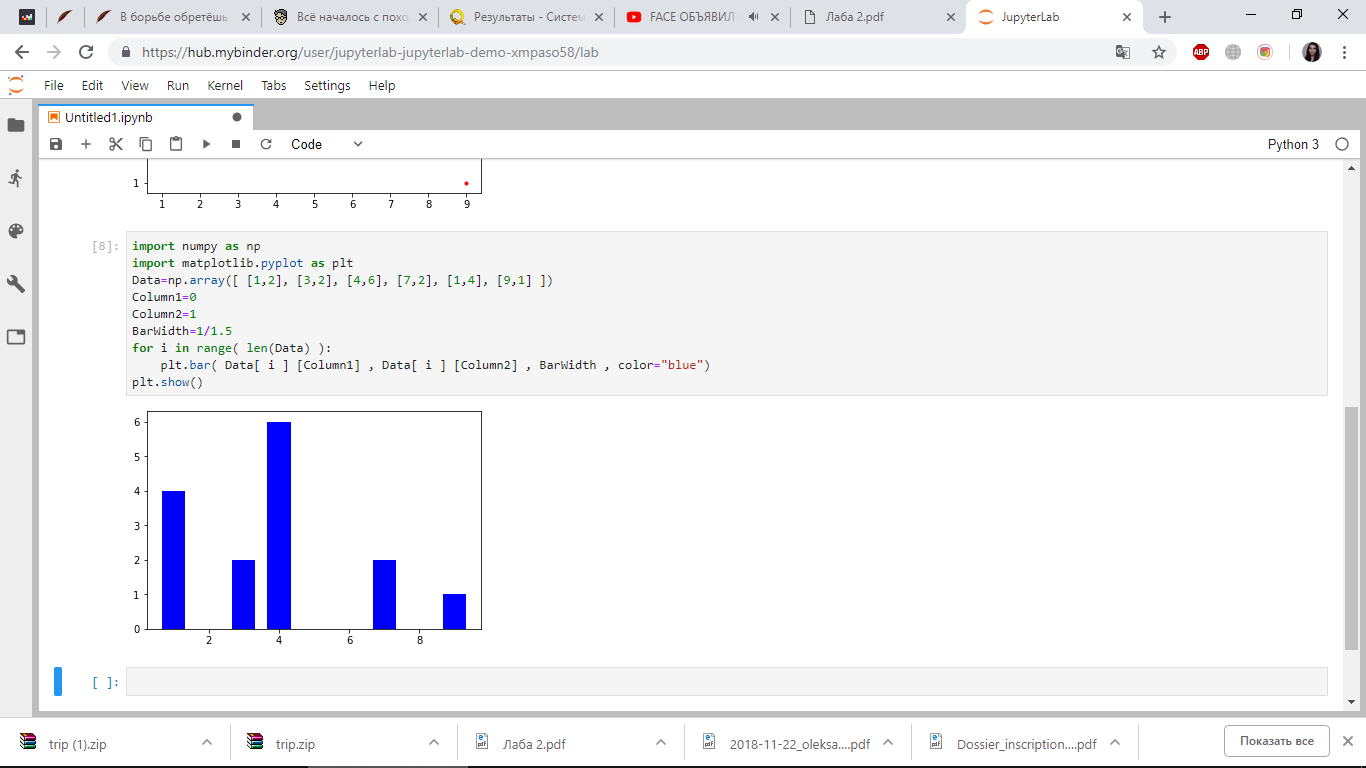
2 14.2500

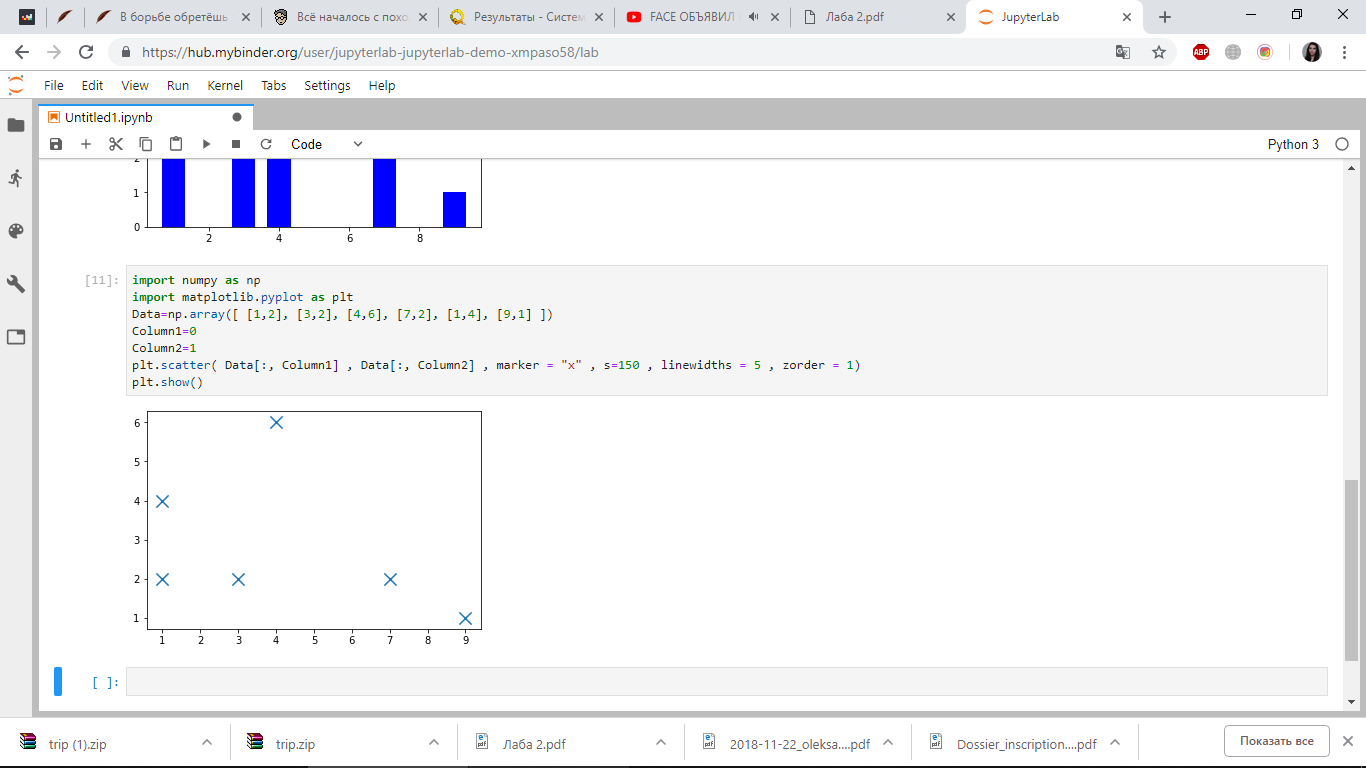
3 8.0500

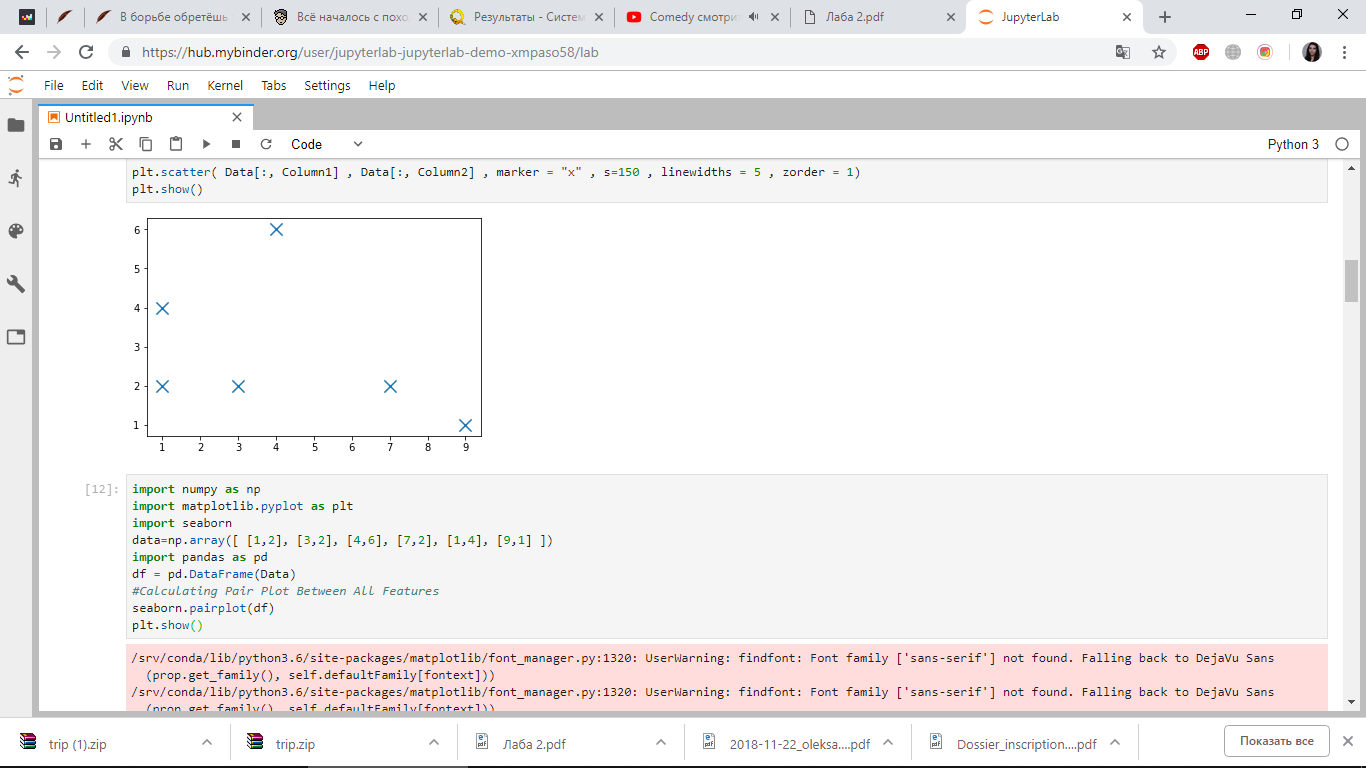
Name: Fare, dtype: float64

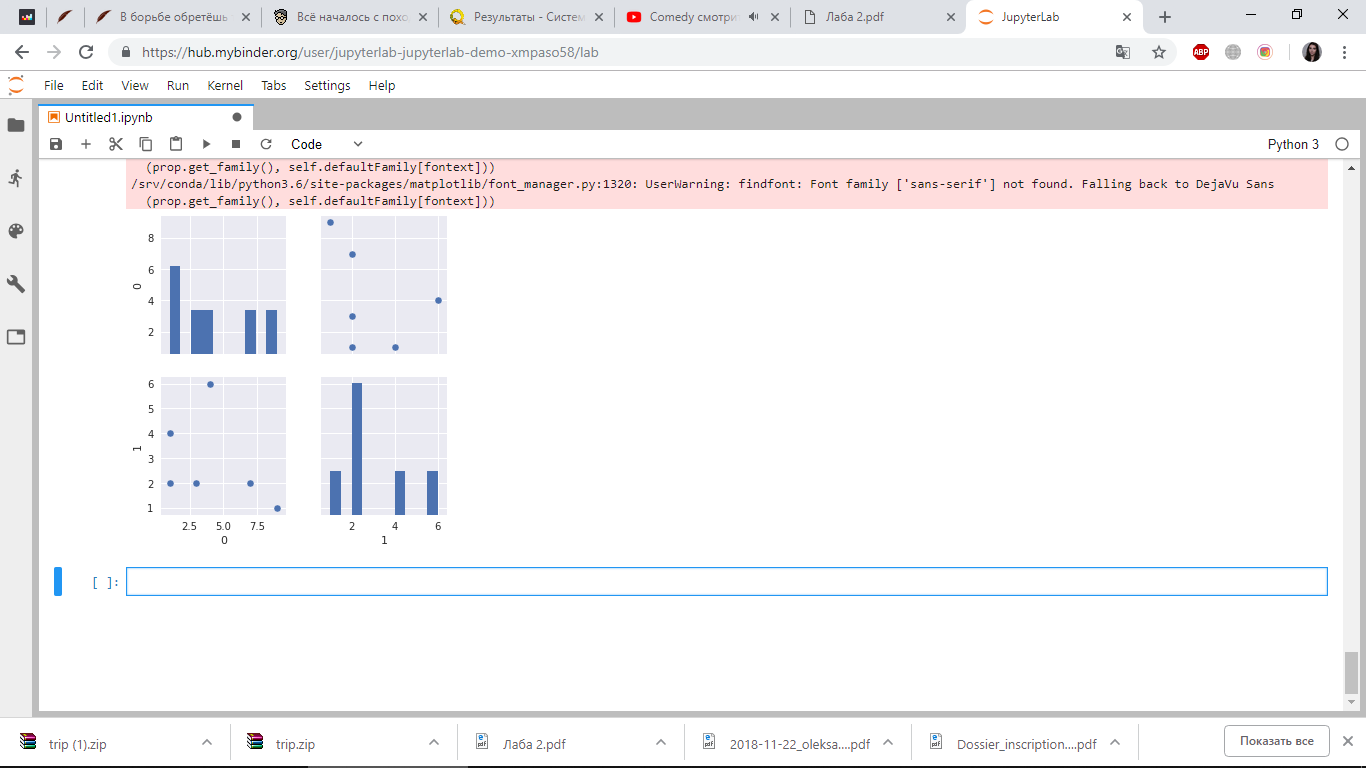
**Лабораторная работа № 2**











Построение Диаграммы с группировкой и условиями