

office online Word | Micro Document | Document | My files - Home Page Download pdf - Jupyter aboutblank

localhost:8888/notebooks/Downloads/pdf.ipynb#

Jupyter pdf Last Checkpoint: 13 minutes ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

```
In [60]: #importing libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import os

In [61]: os.getcwd()
Out[61]: 'C:\\Users\\pc'
```

```
In [1]: path='C:\\Users\\pc\\downloads\\'
data=pd.read_csv(path+'chronickidneydisease.csv')
data.head(10)

-----
NameError                                Traceback (most recent call last)
~\\AppData\\Local\\Temp\\ipykernel_12124\\1579713630.py in <module>
      1 path='C:\\Users\\pc\\downloads\\'
----> 2 data=pd.read_csv(path+'chronickidneydisease.csv')
      3 data.head(10)

NameError: name 'pd' is not defined

In [63]: data.tail(10)
Out[63]:
```

	id	age	bp	sg	al	su	rbc	pc	pcc	ba	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
390	390	52.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	52	6300	5.3	no	no	no	good	no	no	notktd

Type here to search

office online Word | Micro Document | Document | My files - Home Page Downloads pdf - Jupyter aboutblank

localhost:8888/notebooks/Downloads/pdf.ipynb#

Jupyter pdf Last Checkpoint: 13 minutes ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

```
In [63]: data.tail(10)
Out[63]:
```

	id	age	bp	sg	al	su	rbc	pc	pcc	ba	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
390	390	52.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	52	6300	5.3	no	no	no	good	no	no	notktd
391	391	36.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	44	5800	6.3	no	no	no	good	no	no	notktd
392	392	57.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	...	46	6600	5.5	no	no	no	good	no	no	notktd
393	393	43.0	60.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	54	7400	5.4	no	no	no	good	no	no	notktd
394	394	50.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	...	45	9500	4.6	no	no	no	good	no	no	notktd
395	395	55.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	...	47	6700	4.9	no	no	no	good	no	no	notktd
396	396	42.0	70.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	54	7800	6.2	no	no	no	good	no	no	notktd
397	397	12.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	...	49	6600	5.4	no	no	no	good	no	no	notktd
398	398	17.0	60.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	51	7200	5.9	no	no	no	good	no	no	notktd
399	399	58.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	53	6800	6.1	no	no	no	good	no	no	notktd

10 rows x 26 columns

```
In [64]: data.shape
Out[64]: (400, 26)

In [65]: data.columns=['id','age','blood_pressure','specific_gravity','albumin','sugar','red_blood_cells',
'pus_cell','pus_cell_clumps','bacteria','blood_glucose_random',
'blood_urea','serum_creatinine','sodium','potassium','hemoglobin','packed_cell_volume',
'white_blood_cell_count','red_blood_cell_count','hypertension','diabetesmellitus','coronary_artery_disease',
'appetite','pedal_edema','anemia','class']

In [66]: data.columns
Out[66]: Index(['id', 'age', 'blood_pressure', 'specific_gravity', 'albumin', 'sugar',
```

Type here to search

office online Word | Micro Document 16 My files - On Home Page Downloads pdf - Jupyter aboutblank

localhost:8888/notebooks/Downloads/pdf.ipynb#

jupyter pdf Last Checkpoint: 14 minutes ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

```
In [66]: data.columns
Out[66]: Index(['id', 'age', 'blood_pressure', 'specific_gravity', 'albumin', 'sugar',
              'red_blood_cells', 'pus_cell', 'pus_cell_clumps', 'bacteria',
              'blood_glucose_random', 'blood_urea', 'serum_creatinine', 'sodium',
              'potassium', 'hemoglobin', 'packed_cell_volume',
              'white_blood_cell_count', 'red_blood_cell_count', 'hypertension',
              'diabetesmellitus', 'coronary_artery_disease', 'appetite',
              'pedal_edema', 'anemia', 'class'],
              dtype='object')
```

```
In [67]: data.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 400 entries, 0 to 399
Data columns (total 26 columns):
#   Column                Non-Null Count  Dtype
---  -
0   id                    400 non-null   int64
1   age                  391 non-null   float64
2   blood_pressure        388 non-null   float64
3   specific_gravity      353 non-null   float64
4   albumin              354 non-null   float64
5   sugar                351 non-null   float64
6   red_blood_cells       248 non-null   object
7   pus_cell              335 non-null   object
8   pus_cell_clumps       396 non-null   object
9   bacteria              396 non-null   object
10  blood_glucose_random  356 non-null   float64
11  blood_urea            381 non-null   float64
12  serum_creatinine     363 non-null   float64
13  sodium               313 non-null   float64
14  potassium             312 non-null   float64
15  hemoglobin            348 non-null   float64
```

Type here to search 23:46 17-11-2022

office online Word | Micro Document 16 My files - On Home Page Downloads pdf - Jupyter aboutblank

localhost:8888/notebooks/Downloads/pdf.ipynb#

jupyter pdf Last Checkpoint: 16 minutes ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

```
15 hemoglobin            348 non-null   float64
16 packed_cell_volume    330 non-null   object
17 white_blood_cell_count 295 non-null   object
18 red_blood_cell_count  270 non-null   object
19 hypertension          398 non-null   object
20 diabetesmellitus      398 non-null   object
21 coronary_artery_disease 398 non-null   object
22 appetite              399 non-null   object
23 pedal_edema           399 non-null   object
24 anemia                399 non-null   object
25 class                 400 non-null   object
dtypes: float64(11), int64(1), object(14)
memory usage: 81.4+ KB
```

```
In [68]: data.drop(['id'],axis=1,inplace=True)
data
```

```
Out[68]:
```

	age	blood_pressure	specific_gravity	albumin	sugar	red_blood_cells	pus_cell	pus_cell_clumps	bacteria	blood glucose random	packed_cell_volume	white_blo
0	45.0	80.0	1.020	1.0	0.0	NaN	normal	notpresent	notpresent	121.0	...	44
1	7.0	50.0	1.020	4.0	0.0	NaN	normal	notpresent	notpresent	NaN	...	38
2	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent	423.0	...	31
3	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	117.0	...	32
4	51.0	80.0	1.010	2.0	0.0	normal	normal	notpresent	notpresent	106.0	...	35
...	...	...	...	...	...	...	...	...	...	...	...	...
395	55.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	140.0	...	47
396	42.0	70.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	75.0	...	54
397	12.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	100.0	...	49

Type here to search 23:49 17-11-2022

office online Word Document 16 My files - One Home Page Downloads pdf - Jupyter about:blank

localhost:8888/notebooks/Downloads/pdf.ipynb#

jupyter pdf Last Checkpoint: 17 minutes ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

397	12.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	100.0	...	49
398	17.0	60.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	114.0	...	51
399	58.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	131.0	...	53

400 rows x 25 columns

```
In [69]: #target column
data['class'].unique()

Out[69]: array(['ckd', 'ckd\t', 'notckd'], dtype=object)

In [ ]:

In [70]: #rectify target column
data['class']=data['class'].replace('ckd\t','ckd')
data['class'].unique()

Out[70]: array(['ckd', 'notckd'], dtype=object)

In [71]: #fetching categorical column
cat=data.select_dtypes(include=['object']).columns.tolist()
cat

Out[71]: ['red_blood_cells',
'pus_cell',
'pus_cell_clumps',
'bacteria',
'packed_cell_volume',
'white_blood_cell_count',
```

office online Word Document 16 My files - One Home Page Downloads pdf - Jupyter about:blank

localhost:8888/notebooks/Downloads/pdf.ipynb#

jupyter pdf Last Checkpoint: 18 minutes ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

```
'pus_cell',
'pus_cell_clumps',
'bacteria',
'packed_cell_volume',
'white_blood_cell_count',
'red_blood_cell_count',
'hypertension',
'diabetesmellitus',
'coronary_artery_disease',
'appetite',
'pedal_edema',
'anemia',
'class']

In [72]: #removing column which are not categorical
cat.remove('red_blood_cell_count')
cat.remove('packed_cell_volume')
cat.remove('white_blood_cell_count')
cat

Out[72]: ['red_blood_cells',
'pus_cell',
'pus_cell_clumps',
'bacteria',
'hypertension',
'diabetesmellitus',
'coronary_artery_disease',
'appetite',
'pedal_edema',
'anemia',
'class']
```

office online Word | Micro: Document 16: My files - On: Home Page: Downloads: pdf - Jupyter: aboutblank

localhost:8888/notebooks/Downloads/pdf.ipynb#

jupyter pdf Last Checkpoint: 18 minutes ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

```
'class']

In [73]: num=data.select_dtypes(include=['float64']).columns.tolist()#fetch numerical column

In [74]: num.remove('specific_gravity')#remove which are not numerical
num.remove('albumin')
num.remove('sugar')
num

Out[74]: ['age',
'blood_pressure',
'blood_glucose_random',
'blood_urea',
'serum_creatinine',
'sodium',
'potassium',
'hemoglobin']

In [75]: #adding column which is numerical
num.append('red_blood_cell_count')
num.append('packed_cell_volume')
num.append('white_blood_cell_count')
num

Out[75]: ['age',
'blood_pressure',
'blood_glucose_random',
'blood_urea',
'serum_creatinine',
'sodium',
'potassium',
'hemoglobin',
'red_blood_cell_count',
'packed_cell_volume',
'white_blood_cell_count']
```

Type here to search 23:51 17-11-2022

office online Word | Micro: Document 16: My files - On: Home Page: Downloads: pdf - Jupyter: aboutblank

localhost:8888/notebooks/Downloads/pdf.ipynb#

jupyter pdf Last Checkpoint: 18 minutes ago (autosaved) Logout

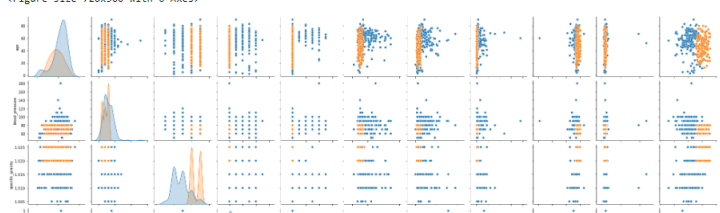
File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

```
'hemoglobin',
'red_blood_cell_count',
'packed_cell_volume',
'white_blood_cell_count']

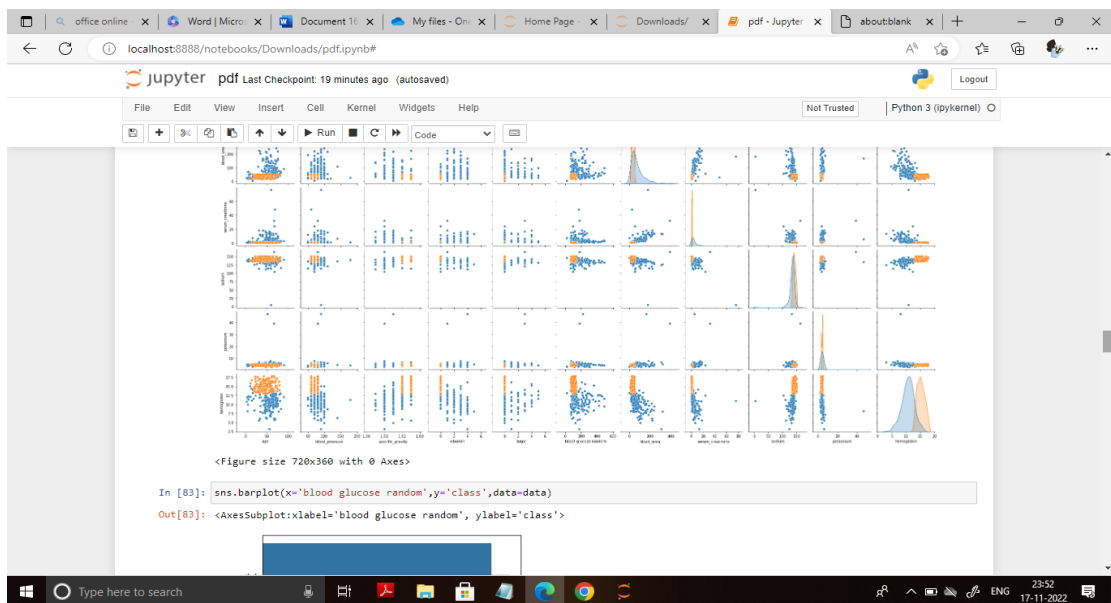
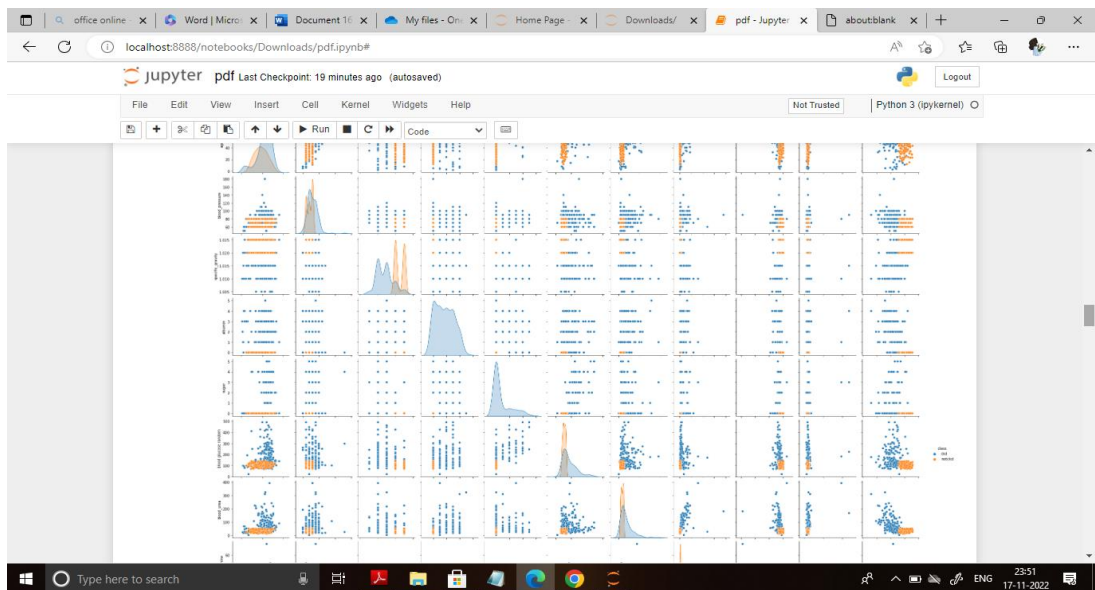
In [82]: sns.pairplot(data,hue='class')
fig=plt.figure(figsize=(10,5))
fig

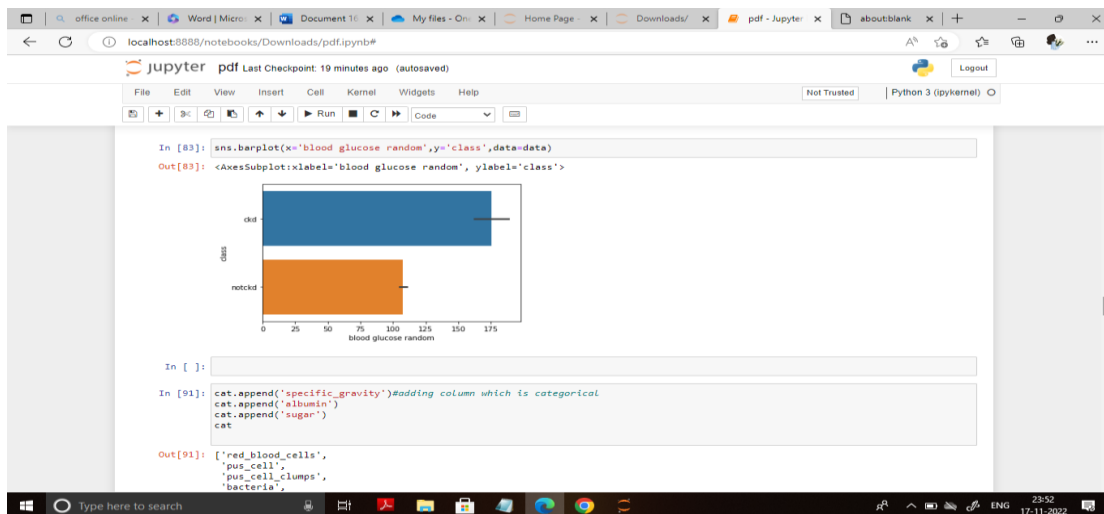
C:\Users\pc\anaconda3\lib\site-packages\seaborn\distributions.py:305: UserWarning: Dataset has 0 variance; skipping density estimate.
warnings.warn(msg, UserWarning)
C:\Users\pc\anaconda3\lib\site-packages\seaborn\distributions.py:305: UserWarning: Dataset has 0 variance; skipping density estimate.
warnings.warn(msg, UserWarning)

Out[82]: <Figure size 720x360 with 0 Axes>
```



Type here to search 23:51 17-11-2022





office online Word | Micro Document 16 My files - On Home Page Downloads pdf - Jupyter aboutblank

localhost:8888/notebooks/Downloads/pdf.ipynb#

jupyter pdf Last Checkpoint: 20 minutes ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

```
In [91]: cat.append('specific_gravity')#adding column which is categorical
cat.append('albumin')
cat.append('sugar')
cat

Out[91]: ['red_blood_cells',
'pus_cell',
'pus_cell_clumps',
'bacteria',
'hypertension',
'diabetesmellitus',
'coronary_artery_disease',
'appetite',
'pedal_edema',
'anemia',
'class',
'specific_gravity',
'albumin',
'sugar']

In [92]: a=data['coronary_artery_disease'].unique()
#b=data['sugar'].unique()
#c=data['albumin'].unique()
#d=data['specific_gravity'].unique()
#e=data['anemia'].unique()
#f=data['pedal_edema'].unique()#
#g=data['appetite'].unique()
#h=data['diabetesmellitus'].unique()
#i=data['bacteria'].unique()
#j=data['hypertension'].unique()#
#k=data['red_blood_cell_count'].unique()
#l=data['pus_cell'].unique()
#m=data['pus_cell_clumps'].unique()
```

office online Word | Micro Document | My files - On Home Page Downloads pdf - Jupyter about:blank

localhost:8888/notebooks/Downloads/pdf.ipynb#

Jupyter pdf Last Checkpoint: 20 minutes ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

```
In [93]: a=data['coronary_artery_disease'].unique()
#b=data['sugar'].unique()
#c=data['albumin'].unique()
#d=data['specific_gravity'].unique()
#e=data['anemia'].unique()
#f=data['pedal_edema'].unique()#
#g=data['appetite'].unique()
#h=data['diabetesmellitus'].unique()
#i=data['bacteria'].unique()
#j=data['hypertension'].unique()#
#k=data['red_blood_cell_count'].unique()
#l=data['pus_cell'].unique()
#m=data['pus_cell_clumps'].unique()

In [93]: a,h
Out[93]: (array(['no', 'yes', '\tno', nan], dtype=object),
          array(['yes', 'no', ' yes', '\tno', '\tyes', nan], dtype=object))

In [94]: #rectifying the categorical column classes
data['coronary_artery_disease']=data['coronary_artery_disease'].replace('\tno','no')
data['coronary_artery_disease'].unique()
Out[94]: array(['no', 'yes', nan], dtype=object)

In [95]: data['diabetesmellitus']=data.diabetesmellitus.replace(to_replace={'yes':'yes','\tyes':'yes','\tno':'no'})
data['diabetesmellitus'].unique()
Out[95]: array(['yes', 'no', ' yes', nan], dtype=object)
```

Type here to search

office online Word | Micro Document | My files - On Home Page Downloads pdf - Jupyter about:blank

localhost:8888/notebooks/Downloads/pdf.ipynb#

Jupyter pdf Last Checkpoint: 21 minutes ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3 (ipykernel)

```
#c=data['albumin'].unique()
#d=data['specific_gravity'].unique()
#e=data['anemia'].unique()
#f=data['pedal_edema'].unique()#
#g=data['appetite'].unique()
#h=data['diabetesmellitus'].unique()
#i=data['bacteria'].unique()
#j=data['hypertension'].unique()#
#k=data['red_blood_cell_count'].unique()
#l=data['pus_cell'].unique()
#m=data['pus_cell_clumps'].unique()

In [93]: a,h
Out[93]: (array(['no', 'yes', '\tno', nan], dtype=object),
          array(['yes', 'no', ' yes', '\tno', '\tyes', nan], dtype=object))

In [94]: #rectifying the categorical column classes
data['coronary_artery_disease']=data['coronary_artery_disease'].replace('\tno','no')
data['coronary_artery_disease'].unique()
Out[94]: array(['no', 'yes', nan], dtype=object)

In [95]: data['diabetesmellitus']=data.diabetesmellitus.replace(to_replace={'yes':'yes','\tyes':'yes','\tno':'no'})
data['diabetesmellitus'].unique()
Out[95]: array(['yes', 'no', ' yes', nan], dtype=object)
```

Type here to search