Assignment 6

Submit:

- Python code
- Evidence that it ran. You can either provide a screenshots of the result in Spyder.

I recommend using x2go for connecting and the Spyder editor (under Programming) for writing Python programs. You can also do this assignment on your home computer if you install all the necessary libraries and a PostgreSQL database. You will need to import the following libraries:

```
import numpy
import pandas
import matplotlib
import psycopg2
```

a) (1 points) Create a tuple
thistuple = ("apple", "banana", "cherry")
and retrieve the element with the index 1

In [28]:
```
import numpy
import pandas
import matplotlib
import psycopg2
```

In [37]:
```
thistuple = ("apple", "banana", "cherry")
print(thistuple[1]);
```

```
banana
```

In [38]:
```
df = pandas.read_csv(r"E:\NDSU\Database\Assignments\Parisa-assignment/train.csv");
print(df)
```

```
      Loan_ID  Gender Married Dependents     Education Self_Employed  \
0    LP001002    Male      No          0      Graduate            No
1    LP001003    Male     Yes          1      Graduate            No
2    LP001005    Male     Yes          0      Graduate           Yes
3    LP001006    Male     Yes          0  Not Graduate            No
4    LP001008    Male      No          0      Graduate            No
..        ...     ...     ...        ...           ...           ...
609  LP002978  Female      No          0      Graduate            No
610  LP002979    Male     Yes         3+      Graduate            No
611  LP002983    Male     Yes          1      Graduate            No
612  LP002984    Male     Yes          2      Graduate            No
613  LP002990  Female      No          0      Graduate           Yes

     ApplicantIncome  CoapplicantIncome  LoanAmount  Loan_Amount_Term  \
```

b) (1 points) Load the data from train.csv making sure to provide the correct path. The print the first 10 rows using the below syntax

```
df = pandas".read_csv("/home/first.last/train.csv")
print(df.head(10))
```

```
In [41]: df = pandas.read_csv(r"E:\NDSU\Database\Assignments\Parisa-assignment/train.csv");
         print(df.head(10))
```

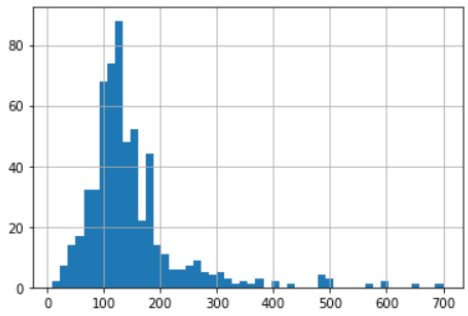
```
    Loan_ID Gender Married Dependents     Education Self_Employed  \
0  LP001002   Male      No          0      Graduate            No
1  LP001003   Male     Yes          1      Graduate            No
2  LP001005   Male     Yes          0      Graduate           Yes
3  LP001006   Male     Yes          0  Not Graduate            No
4  LP001008   Male      No          0      Graduate            No
5  LP001011   Male     Yes          2      Graduate           Yes
6  LP001013   Male     Yes          0  Not Graduate            No
7  LP001014   Male     Yes         3+      Graduate            No
8  LP001018   Male     Yes          2      Graduate            No
9  LP001020   Male     Yes          1      Graduate            No

   ApplicantIncome  CoapplicantIncome  LoanAmount  Loan_Amount_Term  \
0             5849                0.0         NaN             360.0
1             4583             1508.0       128.0             360.0
2             3000                0.0        66.0             360.0
3             2583             2358.0       120.0             360.0
4             6000                0.0       141.0             360.0
5             5417             4196.0       267.0             360.0
6             2333             1516.0        95.0             360.0
7             3036             2504.0       158.0             360.0
8             4006             1526.0       168.0             360.0
9            12841            10968.0       349.0             360.0

   Credit_History Property_Area Loan_Status
0             1.0         Urban           Y
1             1.0         Rural           N
2             1.0         Urban           Y
3             1.0         Urban           Y
4             1.0         Urban           Y
5             1.0         Urban           Y
6             1.0         Urban           Y
7             0.0     Semiurban           N
```

c) (1 points) Show a histogram of LoanAmount using 20 bins. Example of similar type of histogram: `df['ApplicantIncome'].hist(bins=50)`

```
In [40]: df['LoanAmount'].hist(bins=50)
Out[40]: <AxesSubplot:>
```



d) (7points) Import the database interaction library psycopg2. Create a table in the PostgreSQL database that matches the content of the dataframe df (**this can be done using psql**). Insert the content of df into the table you created. The second portion should be done using Python with psycopg2. You can find more documentation at:
https://www.psycopg.org/docs/usage.html and more details in the tutorial at
https://www.postgresqltutorial.com/postgresql-python/

1-As we were asked, I created database table in PSQL and named Loan01.
2-By Copy instruction in Python, I copied the content of train.csv into the Loan01 table.

```
In [29]: #connect to an existing DB
conn = psycopg2.connect(
    host="",
    database="Assignment8",
    user="postgres",
    password="Parisa4653")
```

```
In [30]: cur = conn.cursor()
with open(r"E:\NDSU\Database\Assignments\Parisa-assignment/train.csv") as f:
    next(f)
    cur.copy_from(f, 'loan01', sep=',')

    conn.commit()
```

Then:

In PSQL: Select* from Loan01;

SQL Shell (psql)

```
 loan_status        | boolean              |           |          |
Indexes:
    "loan01_pkey" PRIMARY KEY, btree (loan_id)


Assignment8=# SELECT* from Loan01;
 loan_id  | gender | married | dependents |  education   | self_employed | applicantincome | coapplicantincome | loanamount | loan_amount_term | credit_history | property_area | loan_status
----------+--------+---------+------------+--------------+---------------+-----------------+-------------------+------------+------------------+----------------+---------------+-------------
 LP001002 | Male   | No      | 0          | Graduate     | No            |            5849 |                 0 |            |              360 |              1 | Urban         | t
 LP001003 | Male   | Yes     | 1          | Graduate     | No            |            4583 |              1508 | 128        |              360 |              1 | Rural         | f
 LP001005 | Male   | Yes     | 0          | Graduate     | Yes           |            3000 |                 0 | 66         |              360 |              1 | Urban         | t
 LP001006 | Male   | Yes     | 0          | Not Graduate | No            |            2583 |              2358 | 120        |              360 |              1 | Urban         | t
 LP001008 | Male   | No      | 0          | Graduate     | No            |            6000 |                 0 | 141        |              360 |              1 | Urban         | t
 LP001011 | Male   | Yes     | 2          | Graduate     | Yes           |            5417 |              4196 | 267        |              360 |              1 | Urban         | t
 LP001013 | Male   | Yes     | 0          | Not Graduate | No            |            2333 |              1516 | 95         |              360 |              1 | Urban         | t
 LP001014 | Male   | Yes     | 3+         | Graduate     | No            |            3036 |              2504 | 158        |              360 |              0 | Semiurban     | f
 LP001018 | Male   | Yes     | 2          | Graduate     | No            |            4006 |              1526 | 168        |              360 |              1 | Urban         | t
 LP001020 | Male   | Yes     | 1          | Graduate     | No            |           12841 |             10968 | 349        |              360 |              1 | Semiurban     | f
 LP001024 | Male   | Yes     | 2          | Graduate     | No            |            3200 |               700 | 70         |              360 |              1 | Urban         | t
 LP001027 | Male   | Yes     | 2          | Graduate     |               |            2500 |              1840 | 109        |              360 |              1 | Urban         | t
 LP001028 | Male   | Yes     | 2          | Graduate     | No            |            3073 |              8106 | 200        |              360 |              1 | Urban         | t
 LP001029 | Male   | No      | 0          | Graduate     | No            |            1853 |              2840 | 114        |              360 |              1 | Rural         | f
 LP001030 | Male   | Yes     | 2          | Graduate     | No            |            1299 |              1086 | 17         |              120 |              1 | Urban         | t
 LP001032 | Male   | No      | 0          | Graduate     | No            |            4950 |                 0 | 125        |              360 |              1 | Urban         | t
 LP001034 | Male   | No      | 1          | Not Graduate | No            |            3596 |                 0 | 100        |              240 |                | Urban         | t
 LP001036 | Female | No      | 0          | Graduate     | No            |            3510 |                 0 | 76         |              360 |              0 | Urban         | f
 LP001038 | Male   | Yes     | 0          | Not Graduate | No            |            4887 |                 0 | 133        |              360 |              1 | Rural         | f
 LP001041 | Male   | Yes     | 0          | Graduate     |               |            2600 |              3500 | 115        |                  |              1 | Urban         | t
 LP001043 | Male   | Yes     | 0          | Not Graduate | No            |            7660 |                 0 | 104        |              360 |              0 | Urban         | f
 LP001046 | Male   | Yes     | 1          | Graduate     | No            |            5955 |              5625 | 315        |              360 |              1 | Urban         | t
 LP001047 | Male   | Yes     | 0          | Not Graduate | No            |            2600 |              1911 | 116        |              360 |              0 | Semiurban     | f
 LP001050 |        | Yes     | 2          | Not Graduate | No            |            3365 |              1917 | 112        |              360 |              0 | Rural         | f
 LP001052 | Male   | Yes     | 1          | Graduate     |               |            3717 |              2925 | 151        |              360 |                | Semiurban     | f
 LP001066 | Male   | Yes     | 0          | Graduate     | Yes           |            9560 |                 0 | 191        |              360 |              1 | Semiurban     | t
 LP001068 | Male   | Yes     | 0          | Graduate     | No            |            2799 |              2253 | 122        |              360 |              1 | Semiurban     | t
 LP001073 | Male   | Yes     | 2          | Not Graduate | No            |            4226 |              1040 | 110        |              360 |              1 | Urban         | t
 LP001086 | Male   | No      | 0          | Not Graduate | No            |            1442 |                 0 | 35         |              360 |              1 | Urban         | f
 LP001087 | Female | No      | 2          | Graduate     |               |            3750 |              2083 | 120        |              360 |              1 | Semiurban     | t
 LP001091 | Male   | Yes     | 1          | Graduate     |               |            4166 |              3369 | 201        |              360 |                | Urban         | f
 LP001095 | Male   | No      | 0          | Graduate     | No            |            3167 |                 0 | 74         |              360 |              1 | Urban         | f
 LP001097 | Male   | No      | 1          | Graduate     | Yes           |            4692 |                 0 | 106        |              360 |              1 | Rural         | f
 LP001098 | Male   | Yes     | 0          | Graduate     | No            |            3500 |              1667 | 114        |              360 |              1 | Semiurban     | t
 LP001100 | Male   | No      | 3+         | Graduate     | No            |           12500 |              3000 | 320        |              360 |              1 | Rural         | f
 LP001106 | Male   | Yes     | 0          | Graduate     | No            |            2275 |              2067 |            |              360 |              1 | Urban         | t
 LP001109 | Male   | Yes     | 0          | Graduate     | No            |            1828 |              1330 | 100        |                  |              0 | Urban         | f
 LP001112 | Female | Yes     | 0          | Graduate     | No            |            3667 |              1459 | 144        |              360 |              1 | Semiurban     | t
 LP001114 | Male   | No      | 0          | Graduate     | No            |            4166 |              7210 | 184        |              360 |              1 | Urban         | t
 LP001116 | Male   | No      | 0          | Not Graduate | No            |            3748 |              1668 | 110        |              360 |              1 | Semiurban     | t
 LP001119 | Male   | No      | 0          | Graduate     | No            |            3600 |                 0 | 80         |              360 |              1 | Urban         | f
 LP001120 | Male   | No      | 0          | Graduate     | No            |            1800 |              1213 | 47         |              360 |              1 | Urban         | t
```


Type here to search

-5°C Clear ENG 12:43 AM 4/1/2022

SQL Shell (psql)

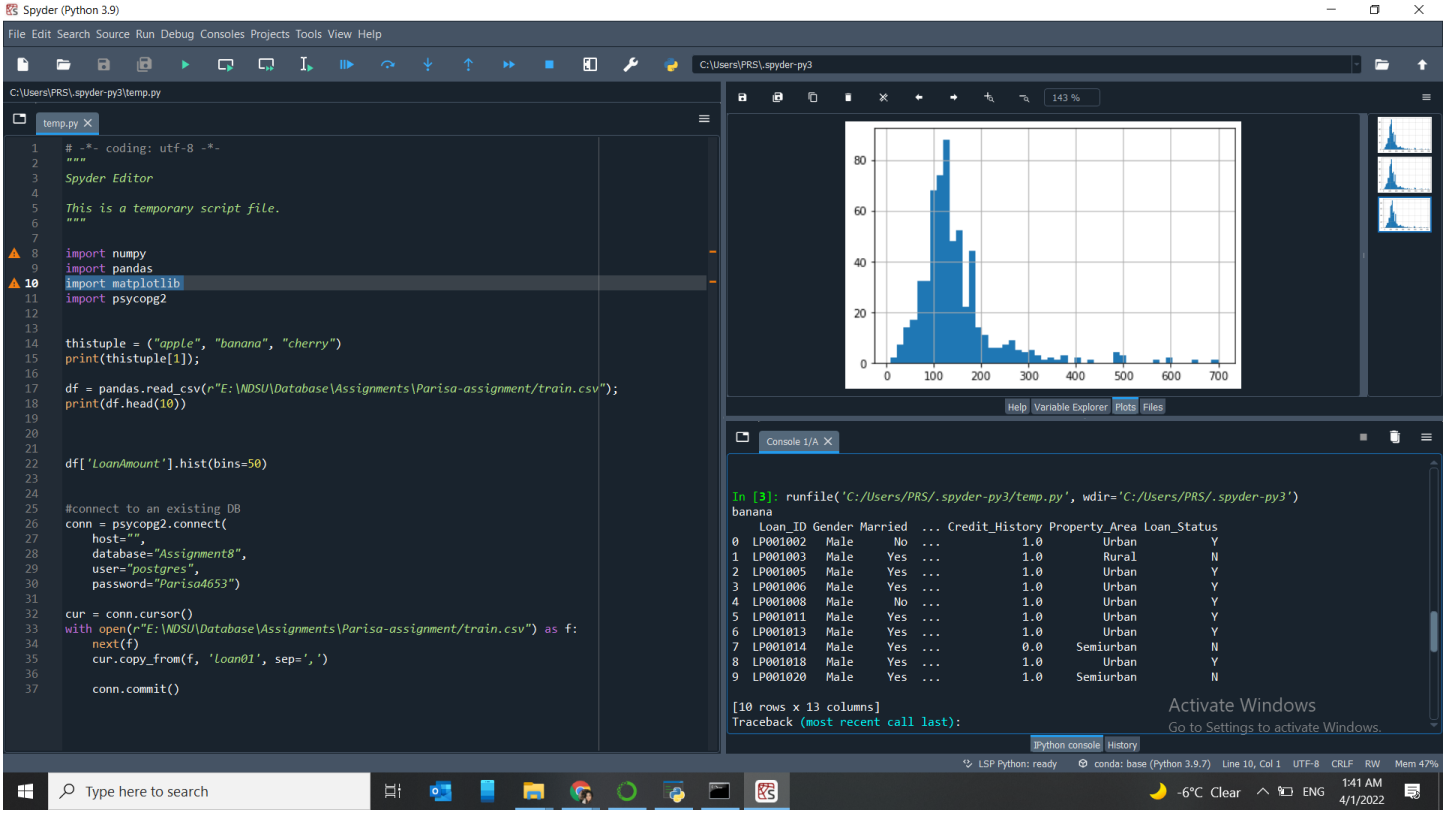
```
 LP002836 | Male   | No      | 0          | Graduate     | No            |            3333 |                 0 | 70         |              360 |              1 | Urban         | t
 LP002837 | Male   | Yes     | 3+         | Graduate     | No            |            3400 |              2500 | 123        |              360 |              0 | Rural         | f
 LP002840 | Female | No      | 0          | Graduate     | No            |            2378 |                 0 | 9          |              360 |              1 | Urban         | f
 LP002841 | Male   | Yes     | 0          | Graduate     | No            |            3166 |              2064 | 104        |              360 |              0 | Urban         | f
 LP002842 | Male   | Yes     | 1          | Graduate     | No            |            3417 |              1750 | 186        |              360 |              1 | Urban         | t
 LP002847 | Male   | Yes     |            | Graduate     | No            |            5116 |              1451 | 165        |              360 |              0 | Urban         | f
 LP002855 | Male   | Yes     | 2          | Graduate     | No            |           16666 |                 0 | 275        |              360 |              1 | Urban         | t
 LP002862 | Male   | Yes     | 2          | Not Graduate | No            |            6125 |              1625 | 187        |              480 |              1 | Semiurban     | f
 LP002863 | Male   | Yes     | 3+         | Graduate     | No            |            6406 |                 0 | 150        |              360 |              1 | Semiurban     | f
 LP002868 | Male   | Yes     | 2          | Graduate     | No            |            3159 |               461 | 108        |               84 |              1 | Urban         | t
 LP002872 |        | Yes     | 0          | Graduate     | No            |            3087 |              2210 | 136        |              360 |              0 | Semiurban     | f
 LP002874 | Male   | No      | 0          | Graduate     | No            |            3229 |              2739 | 110        |              360 |              1 | Urban         | t
 LP002877 | Male   | Yes     | 1          | Graduate     | No            |            1782 |              2232 | 107        |              360 |              1 | Rural         | t
 LP002888 | Male   | No      | 0          | Graduate     |               |            3182 |              2917 | 161        |              360 |              1 | Urban         | t
 LP002892 | Male   | Yes     | 2          | Graduate     | No            |            6540 |                 0 | 205        |              360 |              1 | Semiurban     | t
 LP002893 | Male   | No      | 0          | Graduate     | No            |            1836 |             33837 | 90         |              360 |              1 | Urban         | f
 LP002894 | Female | Yes     | 0          | Graduate     | No            |            3166 |                 0 | 36         |              360 |              1 | Semiurban     | t
 LP002898 | Male   | Yes     | 1          | Graduate     | No            |            1880 |                 0 | 61         |              360 |                | Rural         | f
 LP002911 | Male   | Yes     | 1          | Graduate     | No            |            2787 |              1917 | 146        |              360 |              0 | Rural         | f
 LP002912 | Male   | Yes     | 1          | Graduate     | No            |            4283 |              3000 | 172        |               84 |              1 | Rural         | f
 LP002916 | Male   | Yes     | 0          | Graduate     | No            |            2297 |              1522 | 104        |              360 |              1 | Urban         | t
 LP002917 | Female | No      | 0          | Not Graduate | No            |            2165 |                 0 | 70         |              360 |              1 | Semiurban     | t
 LP002925 |        | No      | 0          | Graduate     | No            |            4750 |                 0 | 94         |              360 |              1 | Semiurban     | t
 LP002926 | Male   | Yes     | 2          | Graduate     | Yes           |            2726 |                 0 | 106        |              360 |              0 | Semiurban     | f
 LP002928 | Male   | Yes     | 0          | Graduate     | No            |            3000 |              3416 | 56         |              180 |              1 | Semiurban     | t
 LP002931 | Male   | Yes     | 2          | Graduate     | Yes           |            6000 |                 0 | 205        |              240 |              1 | Semiurban     | f
 LP002933 |        | No      | 3+         | Graduate     | Yes           |            9357 |                 0 | 292        |              360 |              1 | Semiurban     | t
 LP002936 | Male   | Yes     | 0          | Graduate     | No            |            3859 |              3300 | 142        |              180 |              1 | Rural         | t
 LP002938 | Male   | Yes     | 0          | Graduate     | Yes           |           16120 |                 0 | 260        |              360 |              1 | Urban         | t
 LP002940 | Male   | No      | 0          | Not Graduate | No            |            3833 |                 0 | 110        |              360 |              1 | Rural         | t
 LP002941 | Male   | Yes     | 2          | Not Graduate | Yes           |            6383 |              1000 | 187        |              360 |              1 | Rural         | f
 LP002943 | Male   | No      |            | Graduate     | No            |            2987 |                 0 | 88         |              360 |              0 | Semiurban     | f
 LP002945 | Male   | Yes     | 0          | Graduate     | Yes           |            9963 |                 0 | 180        |              360 |              1 | Rural         | t
 LP002948 | Male   | Yes     | 2          | Graduate     | No            |            5780 |                 0 | 192        |              360 |              1 | Urban         | t
 LP002949 | Female | No      | 3+         | Graduate     |               |             416 |             41667 | 350        |              180 |                | Urban         | f
 LP002950 | Male   | Yes     | 0          | Not Graduate |               |            2894 |              2792 | 155        |              360 |              1 | Rural         | t
 LP002953 | Male   | Yes     | 3+         | Graduate     | No            |            5703 |                 0 | 128        |              360 |              1 | Urban         | t
 LP002958 | Male   | No      | 0          | Graduate     | No            |            3676 |              4301 | 172        |              360 |              1 | Rural         | t
 LP002959 | Female | Yes     | 1          | Graduate     | No            |           12000 |                 0 | 496        |              360 |              1 | Semiurban     | t
 LP002960 | Male   | Yes     | 0          | Not Graduate | No            |            2400 |              3800 |            |              180 |              1 | Urban         | f
 LP002961 | Male   | Yes     | 1          | Graduate     | No            |            3400 |              2500 | 173        |              360 |              1 | Semiurban     | t
 LP002964 | Male   | Yes     | 2          | Not Graduate | No            |            3987 |              1411 | 157        |              360 |              1 | Rural         | t
 LP002974 | Male   | Yes     | 0          | Graduate     | No            |            3232 |              1950 | 108        |              360 |              1 | Rural         | t
 LP002978 | Female | No      | 0          | Graduate     | No            |            2900 |                 0 | 71         |              360 |              1 | Rural         | t
 LP002979 | Male   | Yes     | 3+         | Graduate     | No            |            4106 |                 0 | 40         |              180 |              1 | Rural         | t
 LP002983 | Male   | Yes     | 1          | Graduate     | No            |            8072 |               240 | 253        |              360 |              1 | Urban         | t
 LP002984 | Male   | Yes     | 2          | Graduate     | No            |            7583 |                 0 | 187        |              360 |              1 | Urban         | t
 LP002990 | Female | No      | 0          | Graduate     | Yes           |            4583 |                 0 | 133        |              360 |              0 | Semiurban     | f
(614 rows)
```


Type here to search

-5°C Clear ENG 12:45 AM 4/1/2022

614 records is shown in Loan01 table.

Spyder:


Spyder (Python 3.9)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\PRS\.spyder-py3
C:\Users\PRS\.spyder-py3\temp.py
temp.py
-*- coding: utf-8 -*-
"""
Spyder Editor
This is a temporary script file.
"""
import numpy
import pandas
import matplotlib
import psycopg2
thistuple = ("apple", "banana", "cherry")
print(thistuple[1]);
df = pandas.read_csv(r"E:\NDSU\Database\Assignments\Parisa-assignment/train.csv");
print(df.head(10))
df['LoanAmount'].hist(bins=50)
#connect to an existing DB
conn = psycopg2.connect(
host="",
database="Assignment8",
user="postgres",
password="Parisa4653")
cur = conn.cursor()
with open(r"E:\NDSU\Database\Assignments\Parisa-assignment/train.csv") as f:
next(f)
cur.copy_from(f, 'Loan01', sep=',')
conn.commit()
143 %
Help
Variable Explorer
Plots
Files
Console 1/A
In [3]: runfile('C:/Users/PRS/.spyder-py3/temp.py', wdir='C:/Users/PRS/.spyder-py3')
banana
Loan_ID Gender Married ... Credit_History Property_Area Loan_Status
0 LP001002 Male No ... 1.0 Urban Y
1 LP001003 Male Yes ... 1.0 Rural N
2 LP001005 Male Yes ... 1.0 Urban Y
3 LP001006 Male Yes ... 1.0 Urban Y
4 LP001008 Male No ... 1.0 Urban Y
5 LP001011 Male Yes ... 1.0 Urban Y
6 LP001013 Male Yes ... 1.0 Urban Y
7 LP001014 Male Yes ... 0.0 Semiurban N
8 LP001018 Male Yes ... 1.0 Urban Y
9 LP001020 Male Yes ... 1.0 Semiurban N
[10 rows x 13 columns]
Traceback (most recent call last):
Activate Windows
Go to Settings to activate Windows.
IPython console
History
LSP Python: ready
conda: base (Python 3.9.7)
Line 10, Col 1
UTF-8
CRLF
RW
Mem 47%
Type here to search
-6°C Clear
ENG
1:41 AM
4/1/2022