

IMPORTANT QUESTIONS OF INFORMATICS PRACTICES

I

- (a) Write the difference between array and List.
1. Homogeneous & Heterogeneous.
2. Size cannot change & Size can change
3. Vectorised operation & Scalar operation
4. Less memory size than List
- (b) Ans: - [3 7 5 18]
- (c) Ans: reshape()
- (d)(i) (i)

```
[[ 6  9]
 [19  5]]
```
- (ii)

```
[[19 14]
 [24 12]]
```
- (iii)

```
[[6 9]]
```
- (iv)

```
[[19 14  5]]
```
- (ii) Ans (1) `Arr3=np.hstack([Arr1,Arr2])`
(2) `Arr3=np.concatenate([Arr1,Arr2],axis=1)`
- (e) Write python statement to split the following array **Arr** into three different arrays
(Arr1,Arr2,Arr3) .
- Ans: - `np.split(Arr3,[2,5],axis=1)` or `np.hsplit(Arr3,[2,5])`**
- (ii) Ans: [1. 4. 7. 10.]
- (f) Write a python program to find sum of elements in each row of a two dimensional array of shape (3,3).
import numpy as np
A=np.array([[1,1,2],[3,5,8],[13,21,34]])
print(A)
for i in range(3):
 sum1=0
 for j in range(3):
 sum1=sum1+A[i][j]
 print(sum1)
- (g)

```
[[ 1  1  2]
 [ 9 15 24]
 [26 42 68]]
```
- (h) The covariance tells us that how similar varying two dataset are. A high positive covariance between two dataset shows that that dataset is varying similar. Similarly, high negative number between two dataset means they are very dissimilar.
In python covariance is calculated using the numpy array function `cov()`.
Syntax:
import numpy as np
A=np.array([0,2,4,6,8])
B=np.array([1,3,5,7,9])
cov_array=np.cov(A,B)
print(cov_array)

II

```
(a) d1={"Delhi":10927986,"Kochi":12691836,"Kolkata":4631392}
d2={'Delhi':7216781092,'Kochi':850878126913,"Kolkata":4226784631392}
d3={'Delhi':6.6039457,'Kochi':6.7031476,"Kolkata":9.12849457}
d={'Population':d1,'Avg Income':d2,'Per Capita Income':d3}
df=pd.DataFrame(d)
print(df)
```

(ANY OTHER VALID METHOD ALSO CAN USE)

(b) Write **python statement** for the queries (i) to (iv) and write the output for the python statements (v) to (viii) based on the following DataFrame

(i) Display Minimum Rupees of each item.

```
df.groupby(by='ITEM')['RUPEES'].min()
```

(ii) Display Item and company from the data frame.

```
df[['ITEM','COMPANY']]
```

(iii) Display Item, Company and usd in the descending order of **usd**

```
df[['ITEM','COMPANY','USD']].sort_values(by='USD',ascending=False)
```

(iv) Add new column named 'location' and fill with "Cochin"

```
df['location']=['Cochin','Cochin','Cochin','Cochin']
```

(v) df['COMPANY'].mod()

```
0 LG
```

(vi) df.pivot(index='ITEM',columns='COMPANY',values='USD')

```
COMPANY LG SONY VIDEOCON
```

```
ITEM
```

```
AC 800 50 NaN
```

```
TV 700 NaN 650
```

(vii) df.iloc[0:2,1:3]

```
COMPANY RUPEES
```

```
0 LG 12000
```

```
1 VIDEOCON 10000
```

(viii) df.reindex(columns=['ITEM','USD','VALUE'])

```
ITEM USD VALUE
```

```
0 TV 700 NaN
```

```
1 TV 650 NaN
```

```
2 AC 800 NaN
```

```
3 AC 50 NaN
```

III

(a) (i) Write the output for the python statement based on the DataFrame **cricket** provided below.

	Name	Age	Score
0	Sachin	26	87
1	Dhoni	25	67
2	Virat	25	89
3	Rohit	24	55
4	Shikhir	31	47

(i)df['Age'].quantile([0.25 , 0.5 , 0.75])

```
0.25 25.0
```

```
0.50 25.0
```

```
0.75 26.0
```

(ii)df[['Age','Score']].applymap(np.mean))

```
Age Score
```

```
0 26.0 87.0
```

```
1 25.0 67.0
```

```
2 25.0 89.0
```

```
3 24.0 55.0
```

```
4 31.0 47.0
```

(iii)df[['Age','Score']].apply(np.mean)

```
Age 26.2
```

```
Score 69.0
```

(iv)df.pipe(np.mean)

```
Age 26.2
```

```
Score 69.0
```

- (ii) Write python statement using the above DataFrame **Student** to get the following DataFrame as the output.
`df.pivot(index='Name',columns='Subjects').fillna("")`
- (b) Write a python program to find the Total salary of all employees in the DataFrame **employee** without using any aggregate function.

```
d={'Eno':[1,2,3],'Ename':['Ritesh','Aakash','Manu'],'Salary':[15000,16000,18000]}
df=pd.DataFrame(d)
print(df)
sum1=0
for i in range(len(df)):
    sum1=sum1+df.loc[i,'Salary']
print(sum1)
```

- (c) `fillna(0)`
- (d) `Df.rename({2016:16,2017:17,2018:18,2019:19})`
- (e) Identify the syntax error in the following statements and rewrite the corrected.
- `df1.reindex like(df2) # df1 and df2 are Data Frames.`
 - `del dataframe1['New'] # delete column named New`

(f) Pandas and numpy

(g) Explain the Data Frame function **transform**?

The transform function transforms the aggregate data by repeating summary result for each row in the group and make the result having same shape that of original data
 Eg: `df1.groupby(by='Tutor')['Classes'].transform(np.mean)`

IV

(a) Why the following code is not producing any result ? Why is it giving error?
 Because number of element in A and B are different.

(b) What is pyplot? Is it a Python Library?

Pyplot is one of the interface s of matplotlib library of python. This interface provide simple MATLAB functions that can be used for plotting various types of charts

Pyplot is an interface , not library

(c)

```
import matplotlib.pyplot as pl
App_Name=['Angry Bird','Teen Titan','Marvel Comics','ColorMe','Fun Run','Crazy
Taxi','Igram Pro','Wapp Pro','Matha Formulas']
App_Price_in_Rs=[75,120,190,245,550,55,175,75,140]
Total_Downloads=[197000,209000,414000,196000,272000,311000,213000,455000,2
78000] # common for all
```

(i) A line chart depicting the App Name and prices of the apps

`pl.plot(App_Name,App_Price_in_Rs)`

(ii) A bar chart depicting the App Name and downloads of the apps. The bar should be blue in colour and width should be 0.3

`pl.bar(App_Name,Total_Downloads,color='red',width=0.3)`

(iii) The above chart should have proper titles, x and y axis name and legends.

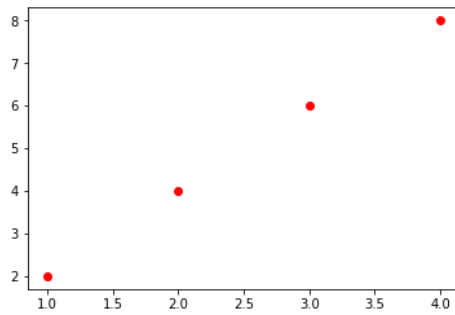
`pl.bar(App_Name,Total_Downloads,color='red',width=0.3,label="Price")`

`pl.legend(loc='left top')`

(iv) A horizontal bar chart to plot App Name and App Price.

`pl.barh(App_Name,App_Price_in_Rs)`

(d) Mr. Sajesh has written the following python code to plot a line chart. What can be the possible line chart?



(e) Python code for plotting the bar chart and its output is given below.
`matplotlib.pyplot.xlim(-3,5)` in this range only one bar is coming.

(f) Weight measurement for 16 small orders of French fries (in grams) are given.

78	72	69	81	63	67	65	75	79	74	71	83	71	79	80	69
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

`weight=[78,72,69,81,63,67,65,75,79,74,71,83,71,79,80,69] #common`

- (i) Create a simple histogram from the above data.
`pl.hist(weight,bins=16)`
- (ii) Create a horizontal histogram from the above data
`pl.hist(weight,bins=16, orientation='horizontal')`
- (iii) Create a step type histogram for the same.
`pl.hist(weight,histtype='step')`
- (iv) Create a notched boxplot for the same.
`pl.boxplot(weight,notch=True)`
- (v) Create a boxplot without the box for the same.
`pl.boxplot(weight,showbox=False)`
- (vi) Show mean in the boxplot for the same.
`pl.boxplot(weight,showmeans=True)`

V

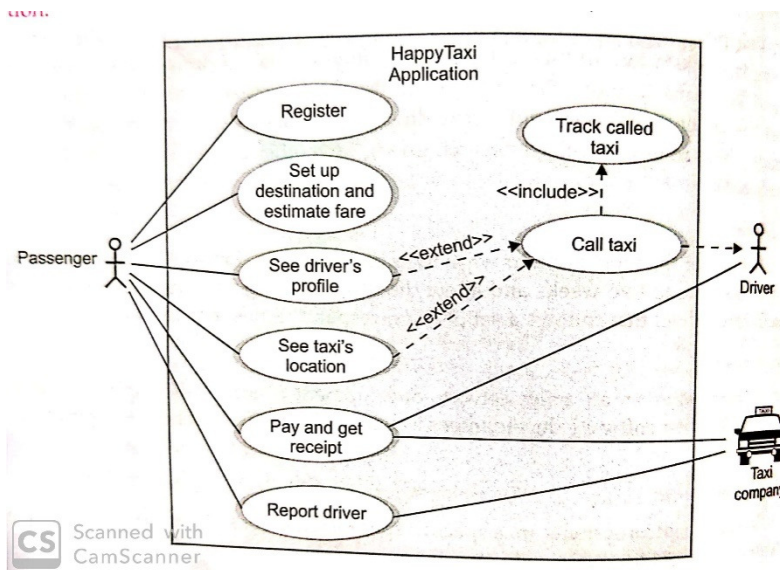
- (a) Waterfall model
- (b) What are the main advantages of spiral model?
 Progressively more complete version of the software gets built from each iteration.
 It is suitable for large, complex and expensive software system.
 It is also suitable for system where reaction to risk required at each evolutionary level.
- (c) Name the delivery model that combines the strength and advantages of waterfall model and evolutionary model? Incremental model
- (d) Verification :- set of activities carried out confirm the software correctly implements the specific functionality.
 Validation means the set of activities that ensures that the software is satisfying the customer's requirement.

- (e) Explain Component based model and list its advantages and disadvantages?
 The component based software engineering is based on the idea to incorporate the and reuse existing software which is already available, in the current software development if feasible or possible.
 Phases:
 1. Component Analysis- Searching and finding out the existing software component suitable for current System.
 2. Required Modification: - Change the requirement so as to reflect in the requirement.
 3. System Design with reuse
 4. Development and Implementation.
 Advantages: reduce amount of software, reduce cost and risk, faster delivery.
 Disadvantage : 1. Requirement compromised.
 2. Some time system does not meet the real need of user.

(f) 'Git' is coming under which type of version control system? -Distributed Version Control system

VI

- (a) Name the **time period** required to create a saleable product in Scrum? - Sprint
 (b) What is the difference between a Commit and update in version control system ?
 Commit is to update the repository
 Update to update the working copy.
 (c) What are the two stereotypes used to define the relation between use cases? <<include>>,<<extend>>
 (d) Draw a use-case for online Taxi Booking app.



- (e) What are the advantages (Utility) of Version control System?
 1. Support partial Edit- Explanation
 2. Enables Multiple people to work simultaneously in same project
 3. Support Multiple Systems. One person can sit on multiple systems while working on same project.
 4. Efficient Integration of work.
 5. Save multiple versions of software.
