Experiment 1

Student Name: Ruchika Raj **UID: 20BCS9285**
Branch: CSE **Section/Group: 615/B**
Semester: 5th **Date of Performance: 04/09/2022**
Subject Name: Machine Learning Lab **Subject Code: 20CSP-317**

1. **Aim/Overview of the practical:** Exploratory Data Analysis on any data set

2. **Source Code:**

```
import pandas as pd
data=pd.read_csv("matches.csv")
print(data)
```

data.dtypes : Return the dtypes in the DataFrame.

data.axes: Return a list representing the axes of the DataFrame.

data.size: Return an int representing the number of elements in this object.

data.shape: Return a tuple representing the dimensionality of the DataFrame.

data.empty: Indicator whether Series/DataFrame is empty.

data.head(): Return the first n rows.(default value of n is 5). **data.head(2)**

data.tail(): Return the last n rows.(default value of n is 5).

data.tail(2)

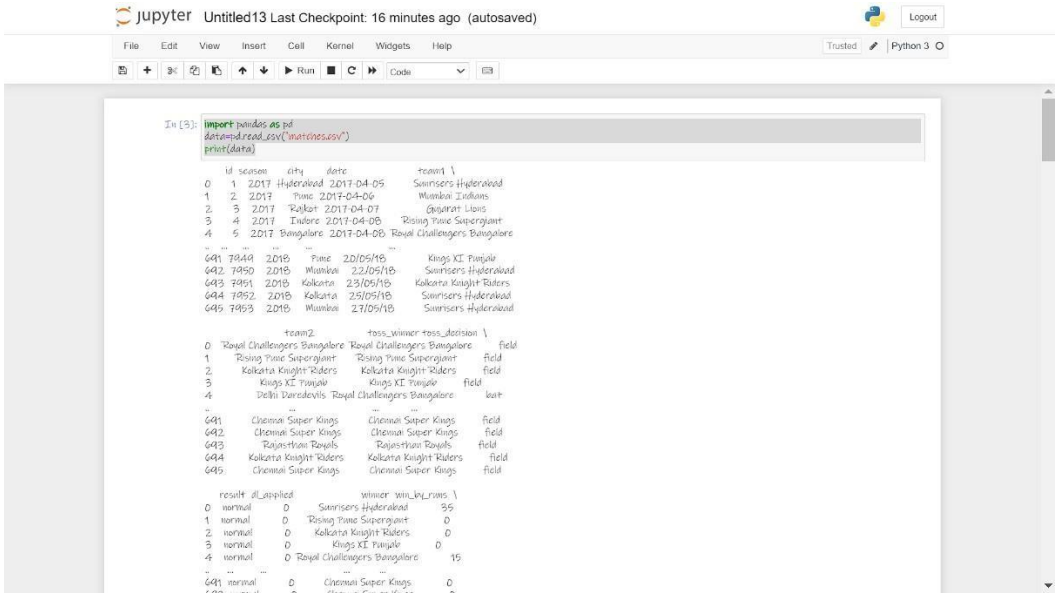
data.count(): Count non-NA cells for each column or row.

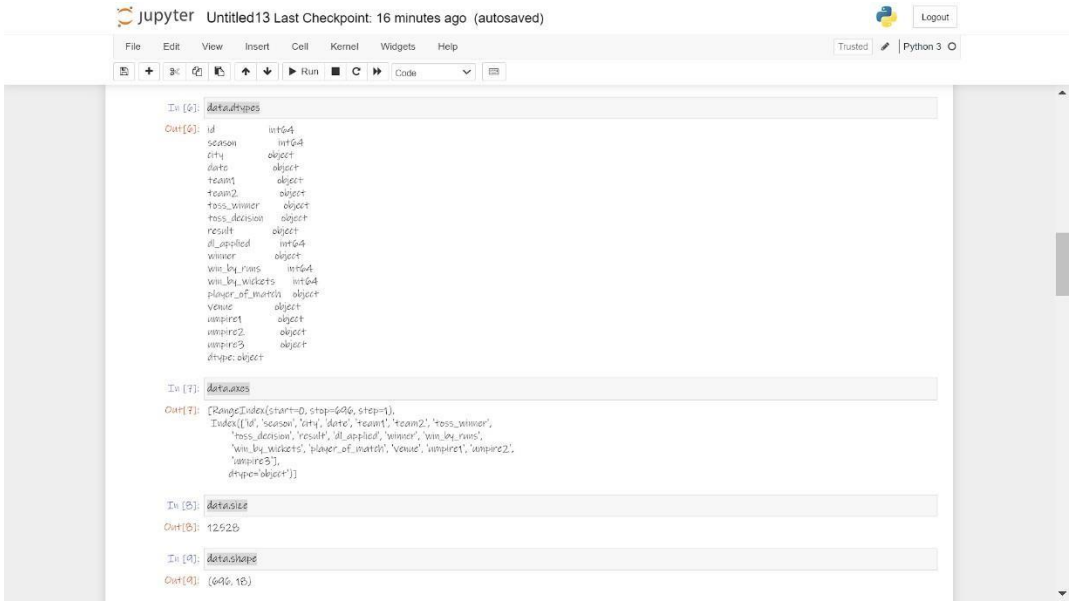
data.sort_values(by='id',ascending=False): Sort by the values along either axis.

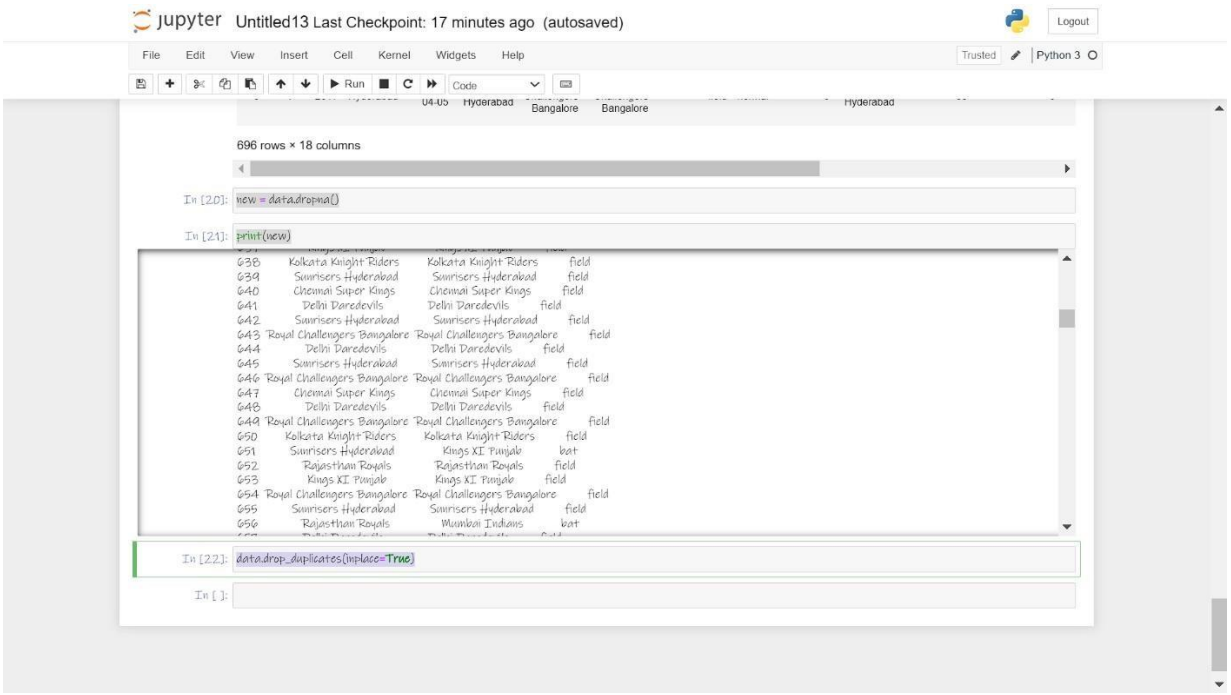
new = data.dropna(): Remove missing values. **print(new)**

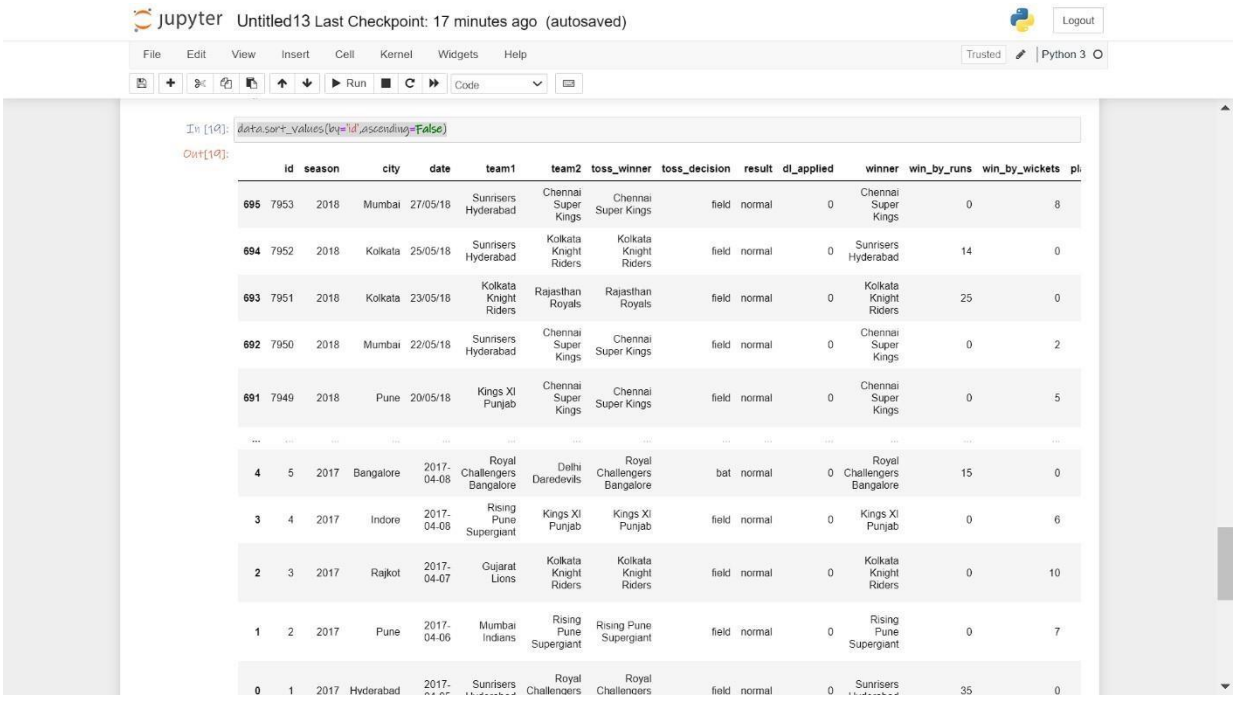
data.drop_duplicates(inplace=True): Return DataFrame with duplicate rows removed.

3. Result/Output


jupyter Untitled13 Last Checkpoint: 16 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted Python 3
Run Code
In [3]:
import pandas as pd
data=pd.read_csv("matches.csv")
print(data)
id season city date team1 \
0 1 2017 Hyderabad 2017-04-05 Sunrisers Hyderabad
1 2 2017 Pune 2017-04-06 Mumbai Indians
2 3 2017 Rajkot 2017-04-07 Gujarat Lions
3 4 2017 Indore 2017-04-08 Rising Pune Supergiant
4 5 2017 Bangalore 2017-04-08 Royal Challengers Bangalore
... ...
691 7948 2019 Pune 20/05/19 Kings XI Punjab
692 7950 2019 Mumbai 22/05/19 Sunrisers Hyderabad
693 7951 2019 Kolkata 23/05/19 Kolkata Knight Riders
694 7952 2019 Kolkata 25/05/19 Sunrisers Hyderabad
695 7953 2019 Mumbai 27/05/19 Sunrisers Hyderabad
team2 toss_winner toss_decision \
0 Royal Challengers Bangalore Royal Challengers Bangalore field
1 Rising Pune Supergiant Rising Pune Supergiant field
2 Kolkata Knight Riders Kolkata Knight Riders field
3 Kings XI Punjab Kings XI Punjab field
4 Delhi Daredevils Royal Challengers Bangalore bat
... ...
691 Chennai Super Kings Chennai Super Kings field
692 Chennai Super Kings Chennai Super Kings field
693 Rajasthan Royals Rajasthan Royals field
694 Kolkata Knight Riders Kolkata Knight Riders field
695 Chennai Super Kings Chennai Super Kings field
result dl_applied winner win_by_runs \
0 normal 0 Sunrisers Hyderabad 35
1 normal 0 Rising Pune Supergiant 0
2 normal 0 Kolkata Knight Riders 0
3 normal 0 Kings XI Punjab 0
4 normal 0 Royal Challengers Bangalore 15
... ...
691 normal 0 Chennai Super Kings 0


jupyter Untitled13 Last Checkpoint: 16 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted Python 3
Run Code
In [6]:
data.dtypes
Out[6]:
id int64
season int64
city object
date object
team1 object
team2 object
toss_winner object
toss_decision object
result object
dl_applied int64
winner object
win_by_runs int64
win_by_wickets int64
player_of_match object
venue object
umpire1 object
umpire2 object
umpire3 object
dtype: object
In [7]:
data.axes
Out[7]:
[RangeIndex(start=0, stop=696, step=1),
Index(['id', 'season', 'city', 'date', 'team1', 'team2', 'toss_winner',
'toss_decision', 'result', 'dl_applied', 'winner', 'win_by_runs',
'win_by_wickets', 'player_of_match', 'venue', 'umpire1', 'umpire2',
'umpire3'],
dtype='object')]
In [8]:
data.size
Out[8]:
12528
In [9]:
data.shape
Out[9]:
(696, 18)





	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	pl
695	7953	2018	Mumbai	27/05/18	Sunrisers Hyderabad	Chennai Super Kings	Chennai Super Kings	field	normal	0	Chennai Super Kings	0	8	
694	7952	2018	Kolkata	25/05/18	Sunrisers Hyderabad	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Sunrisers Hyderabad	14	0	
693	7951	2018	Kolkata	23/05/18	Kolkata Knight Riders	Rajasthan Royals	Rajasthan Royals	field	normal	0	Kolkata Knight Riders	25	0	
692	7950	2018	Mumbai	22/05/18	Sunrisers Hyderabad	Chennai Super Kings	Chennai Super Kings	field	normal	0	Chennai Super Kings	0	2	
691	7949	2018	Pune	20/05/18	Kings XI Punjab	Chennai Super Kings	Chennai Super Kings	field	normal	0	Chennai Super Kings	0	5	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4	5	2017	Bangalore	2017-04-08	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	
3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	0	Kings XI Punjab	0	6	
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	10	
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	

jupyter Untitled13 Last Checkpoint: 17 minutes ago (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

Run | Code

In [14]:
```
data.tail(2)
```

Out[14]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_
694	7952	2018	Kolkata	25/05/18	Sunrisers Hyderabad	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Sunrisers Hyderabad	14	0	Rashi
695	7953	2018	Mumbai	27/05/18	Sunrisers Hyderabad	Chennai Super Kings	Chennai Super Kings	field	normal	0	Chennai Super Kings	0	8	SR \

In [15]:
```
data.count()
```

Out[15]:
```
id              696
season          696
city            689
date            696
team1           696
team2           696
toss_winner     696
toss_decision   696
result          696
dl_applied      696
winner          693
win_by_runs     696
win_by_wickets  696
player_of_match 693
venue           696
umpire1         695
umpire2         695
umpire3          60
dtype: int64
```

In [19]:
```
data.sort_values(by='id',ascending=False)
```

Out[19]:

jupyter Untitled13 Last Checkpoint: 17 minutes ago (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

Run | Code

In [12]:
```
data.head(2)
```

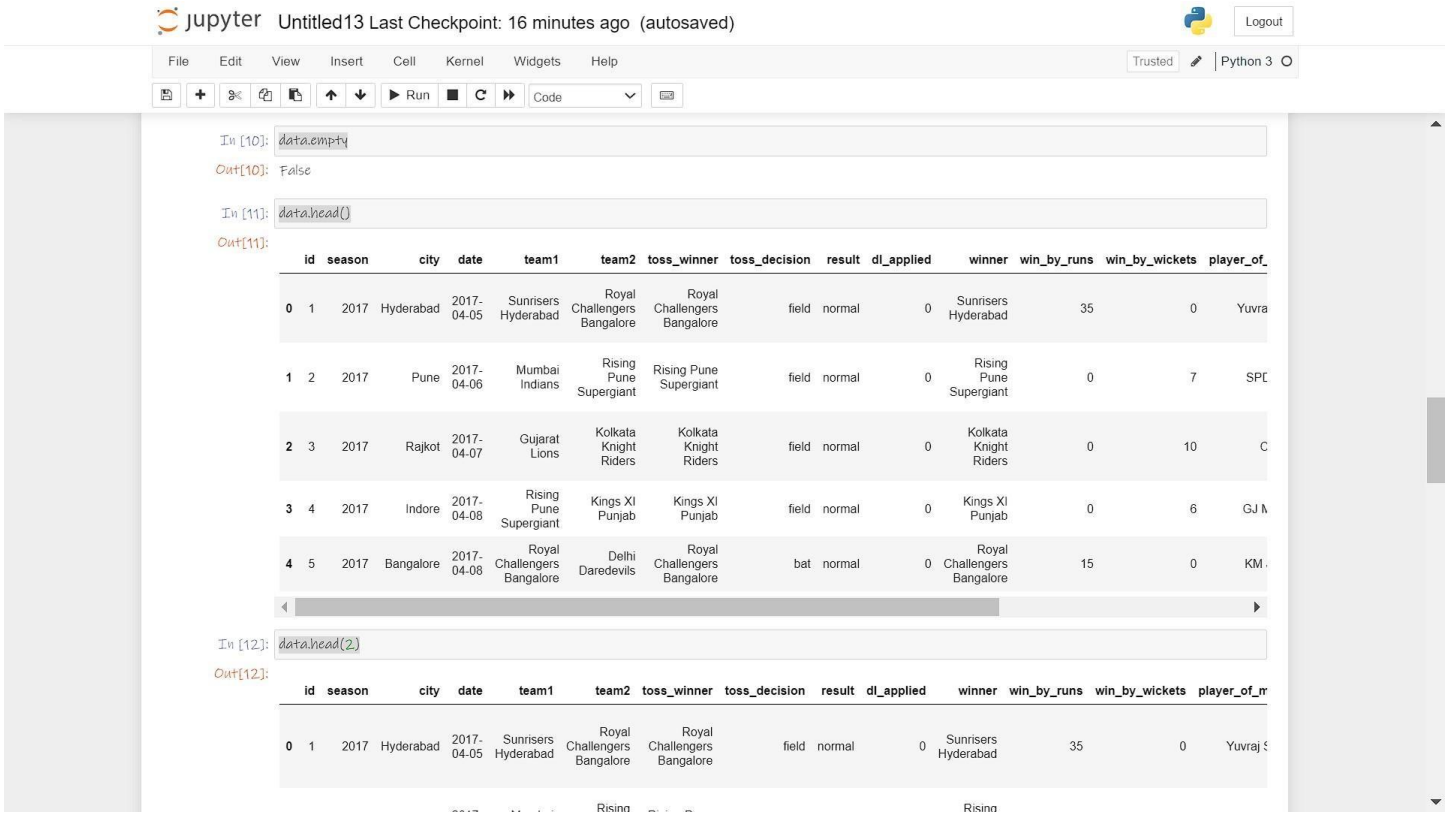
Out[12]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_m
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvraj S
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	SPD S

In [13]:
```
data.tail()
```

Out[13]:

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_o
691	7949	2018	Pune	20/05/18	Kings XI Punjab	Chennai Super Kings	Chennai Super Kings	field	normal	0	Chennai Super Kings	0	5	
692	7950	2018	Mumbai	22/05/18	Sunrisers Hyderabad	Chennai Super Kings	Chennai Super Kings	field	normal	0	Chennai Super Kings	0	2	F d
693	7951	2018	Kolkata	23/05/18	Kolkata Knight Riders	Rajasthan Royals	Rajasthan Royals	field	normal	0	Kolkata Knight Riders	25	0	AD
694	7952	2018	Kolkata	25/05/18	Sunrisers Hyderabad	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Sunrisers Hyderabad	14	0	Ras
695	7953	2018	Mumbai	27/05/18	Sunrisers Hyderabad	Chennai Super Kings	Chennai Super Kings	field	normal	0	Chennai Super Kings	0	8	SR



	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvra
1	2	2017	Pune	2017-04-06	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant	0	7	SPI
2	3	2017	Rajkot	2017-04-07	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders	0	10	C
3	4	2017	Indore	2017-04-08	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	0	Kings XI Punjab	0	6	GJ M
4	5	2017	Bangalore	2017-04-08	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore	15	0	KM

	id	season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_m
0	1	2017	Hyderabad	2017-04-05	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad	35	0	Yuvraj S

Learning outcomes (What I have learnt):

1. **Learned how to use jupyter notebook.**
2. **Learned how to write python programs and its execution.**

3. Learned how to import libraries.

4. Learned how to import dataset in python program using pandas library.

5. Learned how to perform exploratory data analysis on datasets.

Evaluation Grid :

Sr. No.	Parameters	Marks Obtained	Maximum Marks
1.	Student Performance (Conduct of experiment) objectives/Outcomes.		12
2.	Viva Voce		10
3.	Submission of Work Sheet (Record)		8
	Total		30