1. Nba 2021 2022 season activeplayers pandas basics

```
import numpy as np
import pandas as pd
```

Reading of dataset

```
df = pd.read_csv("data/players.csv")
```

 # Les 10 premières données

```
result = df.head(10)
print(result)
```

```
             Name Position            Team  Age  Height  Height_i
Weight  \
0  Juhann Begarin       SG  Boston Celtics   19   6' 5"      6.50
185
1    Jaylen Brown       SG  Boston Celtics   24   6' 6"      6.60
223
2       Kris Dunn       PG  Boston Celtics   27   6' 3"      6.30
205
3  Carsen Edwards       PG  Boston Celtics   23  5' 11"      5.11
200
4      Tacko Fall        C  Boston Celtics   25   7' 5"      7.50
311
5  Bruno Fernando        F  Boston Celtics   23   6' 9"      6.90
240
6      Al Horford        C  Boston Celtics   35   6' 9"      6.90
240
7     Enes Kanter        C  Boston Celtics   29  6' 10"      6.10
250
8     Luke Kornet        C  Boston Celtics   26   7' 2"      7.20
250
9  Romeo Langford       SG  Boston Celtics   21   6' 4"      6.40
216

      College      Salary  Points  Rebounds  Assists
0         NaN         NaN     NaN       NaN      NaN
1  California  26758928.0    24.7       6.0      3.4
2  Providence   5005350.0     1.3       1.5      0.5
3      Purdue   1782621.0     4.0       0.8      0.5
4         UCF         NaN     2.5       2.7      0.2
```

```
5    Maryland   1782621.0     1.5       2.4      0.3
6     Florida  27000000.0    14.2       6.7      3.4
7    Kentucky   1669178.0    11.2      11.0      1.2
8  Vanderbilt         NaN     3.4       2.2      0.8
9     Indiana   3804360.0     3.1       1.9      0.7
```

How many rows are in this dataset?

```
result = len(df.index)
print(result)
```

```
339
```

Average Salary

```
result = df["Salary"].mean()
print(result)
```

```
---------------------------------------------------------------------------
---

KeyError                                  Traceback (most recent call
last)

~/.local/lib/python3.9/site-packages/pandas/core/indexes/base.py in
get_loc(self, key, method, tolerance)
   3620             try:
-> 3621                 return self._engine.get_loc(casted_key)
   3622             except KeyError as err:


~/.local/lib/python3.9/site-packages/pandas/_libs/index.pyx in
pandas._libs.index.IndexEngine.get_loc()


~/.local/lib/python3.9/site-packages/pandas/_libs/index.pyx in
pandas._libs.index.IndexEngine.get_loc()


pandas/_libs/hashtable_class_helper.pxi in
pandas._libs.hashtable.PyObjectHashTable.get_item()


pandas/_libs/hashtable_class_helper.pxi in
pandas._libs.hashtable.PyObjectHashTable.get_item()


KeyError: 'Salary'
```

```
The above exception was the direct cause of the following exception:


KeyError                                  Traceback (most recent call
last)

/tmp/ipykernel_28193/1186721733.py in <module>
----> 1 result = df["Salary"]
      2 print(result)


~/.local/lib/python3.9/site-packages/pandas/core/frame.py in
__getitem__(self, key)
   3503             if self.columns.nlevels > 1:
   3504                 return self._getitem_multilevel(key)
-> 3505             indexer = self.columns.get_loc(key)
   3506             if is_integer(indexer):
   3507                 indexer = [indexer]


~/.local/lib/python3.9/site-packages/pandas/core/indexes/base.py in
get_loc(self, key, method, tolerance)
   3621                 return self._engine.get_loc(casted_key)
   3622             except KeyError as err:
-> 3623                 raise KeyError(key) from err
   3624             except TypeError:
   3625                 # If we have a listlike key,
_check_indexing_error will raise


KeyError: 'Salary'
```

Name of the Player who has Maximum salary

```
result = df[df["Salary"].max() == df["Salary"]]["Name"].iloc[0]
print(result)
```

```
---------------------------------------------------------------------------
---

KeyError                                  Traceback (most recent call
last)

~/.local/lib/python3.9/site-packages/pandas/core/indexes/base.py in
get_loc(self, key, method, tolerance)
   3620             try:
-> 3621                 return self._engine.get_loc(casted_key)
   3622             except KeyError as err:


~/.local/lib/python3.9/site-packages/pandas/_libs/index.pyx in
pandas._libs.index.IndexEngine.get_loc()
```

```
~/.local/lib/python3.9/site-packages/pandas/_libs/index.pyx in
pandas._libs.index.IndexEngine.get_loc()


pandas/_libs/hashtable_class_helper.pxi in
pandas._libs.hashtable.PyObjectHashTable.get_item()


pandas/_libs/hashtable_class_helper.pxi in
pandas._libs.hashtable.PyObjectHashTable.get_item()


KeyError: 'Salary'


The above exception was the direct cause of the following exception:


KeyError                                  Traceback (most recent call
last)

/tmp/ipykernel_28193/2930255266.py in <module>
----> 1 result = df[df["Salary"].max() == df["Salary"]]["Name"].iloc[0]
      2 print(result)


~/.local/lib/python3.9/site-packages/pandas/core/frame.py in
__getitem__(self, key)
   3503             if self.columns.nlevels > 1:
   3504                 return self._getitem_multilevel(key)
-> 3505             indexer = self.columns.get_loc(key)
   3506             if is_integer(indexer):
   3507                 indexer = [indexer]


~/.local/lib/python3.9/site-packages/pandas/core/indexes/base.py in
get_loc(self, key, method, tolerance)
   3621                 return self._engine.get_loc(casted_key)
   3622             except KeyError as err:
-> 3623                 raise KeyError(key) from err
   3624             except TypeError:
   3625                 # If we have a listlike key,
_check_indexing_error will raise


KeyError: 'Salary'
```

 # Which position Kemba Walker does play?

```
result = df[df["Name"] == "Kemba Walker"]["Position"].iloc[0]
print(result)
```

 # Find average salary and age for positions by grouping players by position

```
position_mean = df.groupby(["Position"]).mean()
result = position_mean[["Salary","Age"]]
result = result.round(2)
print(result)
```

 # How many different positions are in the dataset

```
result = df["Position"].nunique()
print(result)
```

 # Calculate how many players are in each position

```
result = df["Position"].value_counts()
print(result)
```

 # Calculate how many people play on each team

```
result = df["Team"].value_counts()
print(result)
```

#Find players with "and" in their name

#------- Option1---------

```
df = df.dropna()
result = df[df["Name"].str.contains("and")]
print(result)
```

 # ------- Option2---------

```
def str_find(name):
    if"and" in name.lower():
        return True
    return False

df.dropna(inplace=True)
result = df[df["Name"].apply(str_find)]
print(result)
```

 # Sorting players by age from smallest to largest, but by score from largest to smallest

 # Sorting players who has max points in his peer

```
df.dropna(inplace=True)
result = df.drop(df.columns[[1,4,5,6,7,8]], axis=1, inplace=True)
result = df.sort_values(by=["Age","Points"],ascending=[True,False])
result = result.drop_duplicates(subset=["Age"])
print(result)
```

```
                Name                    Team  Age  Points  Rebounds
Assists  \
489   Anthony Edwards  Minnesota Timberwolves   20    19.3       4.7
2.9
449   Zion Williamson    New Orleans Pelicans   21    27.0       7.2
3.7
206       Trae Young           Atlanta Hawks   22    25.3       3.9
9.4
16       Jayson Tatum          Boston Celtics   23    26.4       7.4
4.3
338      Devin Booker            Phoenix Suns   24    25.6       4.2
4.3
552  Donovan Mitchell               Utah Jazz   25    26.4       4.4
5.2
113       Zach LaVine           Chicago Bulls   26    27.4       5.0
4.9
69        Joel Embiid     Philadelphia Sixers   27    28.5      10.6
2.8
267      Bradley Beal      Washington Wizards   28    31.3       4.7
4.4
31       Kyrie Irving           Brooklyn Nets   29    26.9       4.8
6.0
313    Kawhi Leonard    Los Angeles Clippers   30    24.8       6.5
5.2
529    Damian Lillard  Portland Trail Blazers   31    28.8       4.2
7.5
25       Kevin Durant           Brooklyn Nets   32    26.9       7.1
5.6
287    Stephen Curry   Golden State Warriors   33    32.0       5.5
5.8
70        Danny Green     Philadelphia Sixers   34     9.5       3.8
1.7
234        Kyle Lowry              Miami Heat   35    17.2       5.4
7.3
348       Chris Paul            Phoenix Suns   36    16.4       4.5
8.9
322   Carmelo Anthony      Los Angeles Lakers   37    13.4       3.1
1.5
232     Udonis Haslem              Miami Heat   41     4.0       1.0
0.0

     total_P+A
489       22.2
449       30.7
206       34.7
16        30.7
338       29.9
```

```
552       31.6
113       32.3
69        31.3
267       35.7
31        32.9
313       30.0
529       36.3
25        32.5
287       37.8
70        11.2
234       24.5
348       25.3
322       14.9
232        4.0
```

Ranking of players who are SG and whose score is higher than 20, in descending order of points

```
# I work on Anaconda as Interpreter
"""
result = df.query('Position == "SG" and Points > 20 or Position == "SG"
and Assists > 5')
result = result.sort_values(by='Points',ascending=False)
print(result)
"""
```

nlargest and nsmallest functions

```
"""
result = df.nlargest(5,'Points')
result = df.nsmallest(5,"Rebounds")
"""
```

Conditional Filtering

```
result = df[(df["Position"] == "SG") | (df["Points"] > 9)].drop(columns=
["Age","Height","College","Salary"]).dropna()
print(result)
```

```
                Name Position            Team  Height_i  Weight  Points
\
1        Jaylen Brown       SG  Boston Celtics       6.6     223    24.7
6          Al Horford        C  Boston Celtics       6.9     240    14.2
7        Enes Kanter         C  Boston Celtics       6.1     250    11.2
9     Romeo Langford       SG  Boston Celtics       6.4     216     3.1
13   Josh Richardson       SG  Boston Celtics       6.5     200    12.1
..               ...      ...             ...       ...     ...     ...
547          Rudy Gay       SF       Utah Jazz       6.8     250    11.4
548       Rudy Gobert        C       Utah Jazz       7.1     258    14.3
```

```
551        Joe Ingles       SG       Utah Jazz       6.8     220    12.1
552  Donovan Mitchell       SG       Utah Jazz       6.1     215    26.4
556     Eric Paschall        F       Utah Jazz       6.6     255     9.5

     Rebounds  Assists  total_P+A
1         6.0      3.4       28.1
6         6.7      3.4       17.6
7        11.0      1.2       12.4
9         1.9      0.7        3.8
13        3.3      2.6       14.7
..        ...      ...        ...
547       4.8      1.4       12.8
548      13.5      1.3       15.6
551       3.6      4.7       16.8
552       4.4      5.2       31.6
556       3.2      1.3       10.8

[260 rows x 9 columns]
```

describe method

```
result = df["Points"].describe()
print(result)
```

Summation of indexes in different columns and sorting

```
df['total_P+A'] = df[['Points','Assists']].sum(axis=1)
df_1 = df.sort_values('total_P+A',ascending=False).dropna().head(20)

df_1['total_P+A+R'] = df_1[['total_P+A','Rebounds']].sum(axis=1)
df_2 = df_1.sort_values('total_P+A+R',ascending=False).dropna().head(20)

df_2 = df_2.drop(columns=
["Height","Height_i","Weight","College","Salary","Position"])
df_1 = df_1.drop(columns=
["Height","Height_i","Weight","College","Salary","Position","total_P+A+R"])


print(df_1)
print(df_2)
```

```
                     Name                    Team  Age  Points
Rebounds  \
287           Stephen Curry   Golden State Warriors   33    32.0
5.5
529          Damian Lillard  Portland Trail Blazers   31    28.8
4.2
267            Bradley Beal      Washington Wizards   28    31.3
4.7
29             James Harden           Brooklyn Nets   32    24.6
7.9
```

```
206                Trae Young           Atlanta Hawks   22    25.3
3.9
336        Russell Westbrook      Los Angeles Lakers   32    22.2
11.5
31               Kyrie Irving           Brooklyn Nets   29    26.9
4.8
25               Kevin Durant           Brooklyn Nets   32    26.9
7.1
357              De'Aaron Fox        Sacremento Kings   23    25.2
3.5
113               Zach LaVine           Chicago Bulls   26    27.4
5.0
552          Donovan Mitchell               Utah Jazz   25    26.4
4.4
69                Joel Embiid     Philadelphia Sixers   27    28.5
10.6
16               Jayson Tatum          Boston Celtics   23    26.4
7.4
449           Zion Williamson    New Orleans Pelicans   21    27.0
7.2
58              Julius Randle         New York Knicks   26    24.1
10.2
313             Kawhi Leonard    Los Angeles Clippers   30    24.8
6.5
338              Devin Booker            Phoenix Suns   24    25.6
4.2
508  Shai Gilgeous-Alexander   Oklahoma City Thunder   23    23.7
4.7
499       Karl-Anthony Towns  Minnesota Timberwolves   25    24.8
10.6
129             Collin Sexton     Cleveland Cavaliers   22    24.3
3.1

     Assists  total_P+A
287      5.8       37.8
529      7.5       36.3
267      4.4       35.7
29      10.8       35.4
206      9.4       34.7
336     11.7       33.9
31       6.0       32.9
25       5.6       32.5
357      7.2       32.4
113      4.9       32.3
552      5.2       31.6
69       2.8       31.3
16       4.3       30.7
449      3.7       30.7
58       6.0       30.1
313      5.2       30.0
338      4.3       29.9
508      5.9       29.6
499      4.5       29.3
129      4.4       28.7
```

```
                        Name                     Team  Age  Points
Rebounds  \
336        Russell Westbrook      Los Angeles Lakers   32    22.2
11.5
29              James Harden           Brooklyn Nets   32    24.6
7.9
287            Stephen Curry   Golden State Warriors   33    32.0
5.5
69               Joel Embiid     Philadelphia Sixers   27    28.5
10.6
529           Damian Lillard  Portland Trail Blazers   31    28.8
4.2
267             Bradley Beal      Washington Wizards   28    31.3
4.7
58             Julius Randle         New York Knicks   26    24.1
10.2
499       Karl-Anthony Towns  Minnesota Timberwolves   25    24.8
10.6
25              Kevin Durant           Brooklyn Nets   32    26.9
7.1
206               Trae Young           Atlanta Hawks   22    25.3
3.9
16              Jayson Tatum          Boston Celtics   23    26.4
7.4
449          Zion Williamson    New Orleans Pelicans   21    27.0
7.2
31              Kyrie Irving           Brooklyn Nets   29    26.9
4.8
113              Zach LaVine           Chicago Bulls   26    27.4
5.0
313            Kawhi Leonard    Los Angeles Clippers   30    24.8
6.5
552         Donovan Mitchell               Utah Jazz   25    26.4
4.4
357             De'Aaron Fox        Sacremento Kings   23    25.2
3.5
508  Shai Gilgeous-Alexander   Oklahoma City Thunder   23    23.7
4.7
338             Devin Booker            Phoenix Suns   24    25.6
4.2
129            Collin Sexton     Cleveland Cavaliers   22    24.3
3.1

     Assists  total_P+A  total_P+A+R
336     11.7       33.9         45.4
29      10.8       35.4         43.3
287      5.8       37.8         43.3
69       2.8       31.3         41.9
529      7.5       36.3         40.5
267      4.4       35.7         40.4
58       6.0       30.1         40.3
499      4.5       29.3         39.9
25       5.6       32.5         39.6
206      9.4       34.7         38.6
```

```
16      4.3       30.7        38.1
449     3.7       30.7        37.9
31      6.0       32.9        37.7
113     4.9       32.3        37.3
313     5.2       30.0        36.5
552     5.2       31.6        36.0
357     7.2       32.4        35.9
508     5.9       29.6        34.3
338     4.3       29.9        34.1
129     4.4       28.7        31.8
```