Mini project (Data science using Python)

Part - 1

We shall now test your skills in using Pandas package. We will be using the games Dataset (https://www.kaggle.com/gutsyrobot/games-data/data) from Kaggle.

Answer each question asked below wrt the games dataset.

Import pandas as pd.

In [1]:
```
import pandas as pd
```

Read games.csv as a dataframe called games.

In [5]:
```
dataframe=pd.read_csv("games-1.csv")
```

Check the head of the DataFrame.

In [6]:
```
dataframe.head()
```

Out[6]:

	id	type	name	yearpublished	minplayers	maxplayers	playingtime	minplaytime	ma
0	1	boardgame	Twilight Struggle	2005	2	2	180	180	
1	2	boardgame	Terra Mystica	2012	2	5	150	60	
2	3	boardgame	Caverna: The Cave Farmers	2013	1	7	210	30	
3	4	boardgame	Through the Ages: A Story of Civilization	2006	2	4	240	240	
4	5	boardgame	Puerto Rico	2002	2	5	150	90	

Use .info() method to find out total number of entries in dataset

In [7]:
```
dataframe.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2665 entries, 0 to 2664
Data columns (total 20 columns):
id                      2665 non-null int64
type                    2665 non-null object
name                    2665 non-null object
yearpublished           2665 non-null int64
minplayers              2665 non-null int64
maxplayers              2665 non-null int64
playingtime             2665 non-null int64
minplaytime             2665 non-null int64
maxplaytime             2665 non-null int64
minage                  2665 non-null int64
users_rated             2665 non-null int64
average_rating          2665 non-null float64
bayes_average_rating    2665 non-null float64
total_owners            2665 non-null int64
total_traders           2665 non-null int64
total_wanters           2665 non-null int64
total_wishers           2665 non-null int64
total_comments          2665 non-null int64
total_weights           2665 non-null int64
average_weight          2665 non-null float64
dtypes: float64(3), int64(15), object(2)
memory usage: 416.5+ KB
```

What is the mean playin time for all games put together ?

In [8]:
```
dataframe['playingtime'].mean()
```

Out[8]:
```
105.03489681050657
```

What is the highest number of comments received for a game?

In [9]:
```
dataframe['total_comments'].max()
```

Out[9]:
```
11798
```

What is the name of the game with id 1500?

In [83]:
```
dataframe.loc[dataframe.id==1500,['id','name']]
```

Out[83]:

	id	name
1499	1500	El Grande: Knig & Intrigant

And which year was it published?

In [84]:
```
dataframe.loc[dataframe.id==1500,['name','yearpublished']]
```

Out[84]:

	name	yearpublished
1499	El Grande: Knig & Intrigant	1997

Which game has received highest number of comments?

In [43]:
```
dataframe[dataframe['total_comments']==dataframe['total_comments'].max()]['name']
```

Out[43]:
```
165    Catan
Name: name, dtype: object
```

Which games have received least number of comments?

In [71]:
```
dataframe[dataframe['total_comments']==dataframe['total_comments'].min()]['name']
```

Out[71]:
```
1572      Fire On The Suns: Tactical Command Fleet Book 1
1584                          Hossa! Seefahrer-Erweiterung
1665    Worlds of Heroes & Tyrants: Hell Card Expansio...
1750                  Aura Battler Dunbine: Wing Caliver
1751                 Mobile Suit Z Gundam: Gate of Zedan
1752                      Mobile Suit Gundam: White Base
1753                                      L-Gaim Mark II
1768                  Lemlican Series 3: Arms Collection
1769                Lemlican Series 2: Lemlican Monsters
1770          Wizardry Card Game Wiz Ball Expansion Kit
1772                                      Gal Master 2
1776                           Super Nova: Event Horizon
1785    Aventuras Hericas: A Revelao da Princesa / O E...
1811                     Speed Circuit Tournament Tracks
1889                Sturmovik, Clash of Eagles Expansion
1933                        Forte Trivia Cards Volume II
1934                                            Zeppelin
1961              Love, Sex and Romance Trivia Card Set
1966                                Trivia Adventure Plus
1989                                              Mayhem
2065                Entrepreneur's Accessory to Monopoly
2092                                             Crusade
2098    Starfleet Wars: Observer's Directory & Identif...
2100                 Suomi: A module for Clash of Eagles
2142                                      Bar-barians!
2169                                  Carnage: Skullbrawl
2173       Horsepower: 5000  Expansion Set 1  World Rails
2183                                           Alexander
2188                         Mobile Suit Gundam ZZ: Core 3
2191          Renegade Legion: Interceptor  The Fire Eagles
                              ...
2335                         Napoleonic Scenarios Volume 2
2345    No Pasaran!: La Quinta del Biberon  Balaguer 1938
2383                         Cranium New York Booster Pack
2394                                   Floaties & Sinkies!
2397                               Hossa!: Arbeiterlieder
2429                                           Meck Wars
2430                        Train Raider: Europe Expansion
2448                                 Tank War: Expansion A
2455    Avanti Savoia: Intelligence Handbook on Italia...
2545                  History of War: "Fall Gelb" Edition
2547                        Vampire Wars: The Antagonists
2557                                            Krebiz-4
2606                               Overlords Monsters Packs
2607           Piglings Revenge, PigWars Expansion Module #1
2621    Mad Scientist University Course Packet: Anthro...
2622    Mad Scientist University Course Packet: Chemistry
2623    Mad Scientist University Course Packet: Astronomy
2624    Mad Scientist University Course Packet: Computers
2625    Mad Scientist University Course Packet: Dorm Life
2626    Mad Scientist University Course Packet: Greek ...
2627    Mad Scientist University Course Packet: Physic...
2628     Mad Scientist University Course Packet: Medicine
2629      Mad Scientist University Course Packet: Physics
```

```
2631      Mad Scientist University Course Packet: Theater
2632    Mad Scientist University Course Packet: Winter...
2633      Mad Scientist University Course Packet: Zoology
2635      Mad Scientist University Course Packet: PSI PHI
2636    Mad Scientist University Course Packet: Indepe...
2637    Mad Scientist University Course Packet: Home E...
2649                     La die Kirche ins Dorf Erweiterung
Name: name, Length: 65, dtype: object
```

What was the average minage of all games per game "type"? (boardgame & boardgameexpansion)

In [48]:
```
dataframe.groupby(['type'])['minage'].mean()
```

Out[48]:
```
type
boardgame             10.718499
boardgameexpansion    10.633419
Name: minage, dtype: float64
```

In [47]:
```
dataframe.groupby(['type','name'])['minage'].mean()
```

Out[47]:

type	name	
boardgame	10 Days in Africa	10
	10 Days in Asia	10
	10 Days in Europe	10
	10 Days in the USA	10
	1775: Rebellion	10
	1812: The Invasion of Canada	10
	1830: Railways & Robber Barons	14
	1835	16
	1846	14
	1856	14
	1860: Railways on the Isle of Wight	13
	1861: The Railways of the Russian Empire	12
	1870	14
	1889	14
	18AL	14
	1944: Race to the Rhine	14
	1960: The Making of the President	12
	1989: Dawn of Freedom	12
	1st & Goal	13
	2 de Mayo	12
	20th Century	13
	51st State	10
	6 nimmt!	8
	7 Ages	14
	7 Wonders	10
	A Brief History of the World	10

```
                        A Castle for All Seasons
                    10
                        A Distant Plain
                    12
                        A Few Acres of Snow
                    14
                        A Game of Thrones (first edition)
                    12

                    ..
boardgameexpansion  World At War Expansion
                    12
                        World War II Expansion 2
                    12
                        World War II Expansion 3: The Battle of Midway
                    12
                        World War II The Expansion
                    12
                        World of Warcraft: The Boardgame  Shadow of War
                    12
                        Worlds of Heroes & Tyrants Card Expansion #1: Frozen in Tim
e                    8
                        Worlds of Heroes & Tyrants Card Expansion #2: Secrets of Th
e Cataclysm          8
                        Worlds of Heroes & Tyrants: Hell
                     8
                        Worlds of Heroes & Tyrants: Hell Card Expansion #1  Alone
                     8
                        Worlds of Heroes & Tyrants: Hell Card Expansion #2  Anger M
anagement            8
                        Worlds of Heroes & Tyrants: Hell Card Expansion #3  The Maz
e                    8
                        Worlds of Heroes & Tyrants: Vortex
                     8
                        XIG 4E: Air Pathway
                     8
                        XIG 4E: Fire Pathway
                     0
                        Xenophon
                    12
                        Yanks: ASL Module 3
                    12
                        Ys+
                    10
                        ZRCV: Flying Flat-Top
                    12
                        Za Stalina: Handbook on Soviet Armoured and Cavalry forces
                    12
                        Zargo's Lords 2: Expansion Pack
                    12
                        Zauberschwert & Drachenei: Poison Expansion
                    12
                        Zeppelin
                    12
                        Zombie Plague: Canisters
                    10
                        Zombie Plague: Crawlers
```

```
10
     Zombie Plague: The Cellar
10
     Zombie Plague: Twitchers
10
     Zombies!!! 2: Zombie Corps(e)
16
     Zombies!!! 3.5:  Not Dead Yet
12
     Zombies!!! 3:  Mall Walkers
16
     Zombies!!! 5: School's Out Forever
16
Name: minage, Length: 2658, dtype: int64
```

How many unique games are there in the dataset?

In [75]:
```
dataframe['name'].nunique()
```

Out[75]:
```
2657
```

How many boardgames and boardgameexpansions are there in the dataset?

In [63]:
```
dataframe.groupby(['type'])['type'].count()
```

Out[63]:
```
type
boardgame             1492
boardgameexpansion    1173
Name: type, dtype: int64
```

Is there a correlation between playing time and total comments for the games? - Use the .corr() function

In [64]:
```
dataframe['playingtime'].corr(dataframe['total_comments'])
```

Out[64]:
```
-0.010366606124068173
```

Part 2

Inferential Statistical Analysis

A Brand new Gaming Design Company has entered the Market in year 2015. They are interested in Analyzing the Amount of Time Gamers spend for each Game Type and check if the players spend equal Time on both Types.

Analyze the same using a suitable Two Sample Test at 95% Confidence.

The following points are necessarily to be adhered.

- Clean Data

- Transform Data(as required)
- Visualize Data(Draw as many patterns, not limited to Histograms & Boxplots only) - Minimum Requirement of 5 different Visualizations apart from Histograms & Boxplots to be created.
- Model the Data with the correct hypothesis statement
- Conclude on your results based on the outcome of experiment
- Calculate power of test and interpret the same

The END