

Московский государственный технический университет им. Н.Э.Баумана  
Факультет «Информатика и системы управления»  
Кафедра «Системы обработки информации и управления»



## **“Методы машинного обучения”**

### **“ Обработка пропусков в данных, кодирование категориальных признаков, масштабирование данных”**

#### **Лабораторная работа № 3**

ИСПОЛНИТЕЛЬ:

Студент группы ИУ5-21М

Гузилов А.В. \_\_\_\_\_

ПРЕПОДАВАТЕЛЬ:

Гапанюк Ю.Е. \_\_\_\_\_

Москва      2019

---

**Цель лабораторной работы:** изучение способов предварительной обработки данных для дальнейшего формирования моделей.

**Задание:**

1. Выбрать набор данных (датасет), содержащий категориальные признаки и пропуски в данных. Для выполнения следующих пунктов можно использовать несколько различных наборов данных (один для обработки пропусков, другой для категориальных признаков и т.д.)
2. Для выбранного датасета (датасетов) на основе материалов лекции решить следующие задачи:
  - обработку пропусков в данных;
  - кодирование категориальных признаков;
  - масштабирование данных.

# lab3

April 20, 2019

```
In [3]: import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
sns.set(style="ticks")
```

```
In [6]: data = pd.read_csv('marvel-wikia-data.csv', sep=",")
data.head(17000)
```

```
Out[6]:
```

	page_id	name \
0	1678	Spider-Man (Peter Parker)
1	7139	Captain America (Steven Rogers)
2	64786	Wolverine (James \"Logan\" Howlett)
3	1868	Iron Man (Anthony \"Tony\" Stark)
4	2460	Thor (Thor Odinson)
5	2458	Benjamin Grimm (Earth-616)
6	2166	Reed Richards (Earth-616)
7	1833	Hulk (Robert Bruce Banner)
8	29481	Scott Summers (Earth-616)
9	1837	Jonathan Storm (Earth-616)
10	15725	Henry McCoy (Earth-616)
11	1863	Susan Storm (Earth-616)
12	7823	Namor McKenzie (Earth-616)
13	2614	Oro Munroe (Earth-616)
14	1803	Clinton Barton (Earth-616)
15	1396	Matthew Murdock (Earth-616)
16	55534	Stephen Strange (Earth-616)
17	1978	Mary Jane Watson (Earth-616)
18	1872	John Jonah Jameson (Earth-616)
19	35350	Robert Drake (Earth-616)
20	1557	Henry Pym (Earth-616)
21	65255	Charles Xavier (Earth-616)
22	1073	Warren Worthington III (Earth-616)
23	1346	Piotr Rasputin (Earth-616)
24	2512	Wanda Maximoff (Earth-616)
25	1671	Nicholas Fury (Earth-616)
26	1976	Janet van Dyne (Earth-616)

27	65219	Jean Grey (Earth-616)
28	6545	Natalia Romanova (Earth-616)
29	2223	Kurt Wagner (Earth-616)
...	...	...
16346	42106	Thiazzi (Earth-616)
16347	462671	TOR\test
16348	725312	Toxin (Luminals) (Earth-616)
16349	507239	Urizen Ul'var (Earth-616)
16350	40061	Valka (Earth-616)
16351	505759	Vayor (Earth-616)
16352	16569	Viridian (Earth-616)
16353	16296	William Burke (Earth-616)
16354	120833	William Falsworth (Earth-616)
16355	693077	William Shakespeare (Earth-616)
16356	473710	Yamata no Oricchi (Earth-616)
16357	670512	Zero G. Priestly (Legion Personality) (Earth-616)
16358	12132	Zora Loftus (Earth-616)
16359	727035	Agar (Earth-616)
16360	703546	Ana (Natasha Romanoff's neighbor) (Earth-616)
16361	713903	Dante's mother (Earth-616)
16362	508693	Farbauti (Earth-616)
16363	694577	Finch (Kate Bishop's neighbor) (Earth-616)
16364	655292	Jack O'Lantern (Impostor) (Earth-616)
16365	684262	K'thol (Earth-616)
16366	643435	Karen (Hijack's girlfriend) (Earth-616)
16367	694583	Marcus (Kate Bishop's neighbor) (Earth-616)
16368	703892	Marcy (Offer's employee) (Earth-616)
16369	660799	Melanie Kapoor (Earth-616)
16370	674414	Phoenix's Shadow (Earth-616)
16371	657508	Ru'ach (Earth-616)
16372	665474	Thane (Thanos' son) (Earth-616)
16373	695217	Tinkerer (Skrull) (Earth-616)
16374	708811	TK421 (Spiderling) (Earth-616)
16375	673702	Yologarch (Earth-616)

	urlslug	ID \
0	\/Spider-Man_(Peter_Parker)	Secret Identity
1	\/Captain_America_(Steven_Rogers)	Public Identity
2	\/Wolverine_(James_%22Logan%22_Howlett)	Public Identity
3	\/Iron_Man_(Anthony_%22Tony%22_Stark)	Public Identity
4	\/Thor_(Thor_Odinson)	No Dual Identity
5	\/Benjamin_Grimm_(Earth-616)	Public Identity
6	\/Reed_Richards_(Earth-616)	Public Identity
7	\/Hulk_(Robert_Bruce_Banner)	Public Identity
8	\/Scott_Summers_(Earth-616)	Public Identity
9	\/Jonathan_Storm_(Earth-616)	Public Identity
10	\/Henry_McCoy_(Earth-616)	Public Identity
11	\/Susan_Storm_(Earth-616)	Public Identity

12	\\Namor_McKenzie_(Earth-616)	No Dual Identity
13	\\Ororo_Munroe_(Earth-616)	Public Identity
14	\\Clinton_Barton_(Earth-616)	Public Identity
15	\\Matthew_Murdock_(Earth-616)	Public Identity
16	\\Stephen_Strange_(Earth-616)	Public Identity
17	\\Mary_Jane_Watson_(Earth-616)	No Dual Identity
18	\\John_Jonah_Jameson_(Earth-616)	No Dual Identity
19	\\Robert_Drake_(Earth-616)	Secret Identity
20	\\Henry_Pym_(Earth-616)	Public Identity
21	\\Charles_Xavier_(Earth-616)	Public Identity
22	\\Warren_Worthington_III_(Earth-616)	Public Identity
23	\\Piotr_Rasputin_(Earth-616)	Secret Identity
24	\\Wanda_Maximoff_(Earth-616)	Public Identity
25	\\Nicholas_Fury_(Earth-616)	No Dual Identity
26	\\Janet_van_Dyne_(Earth-616)	Public Identity
27	\\Jean_Grey_(Earth-616)	Public Identity
28	\\Natalia_Romanova_(Earth-616)	Public Identity
29	\\Kurt_Wagner_(Earth-616)	Secret Identity
...	...	...
16346	\\Thiazzi_(Earth-616)	NaN
16347	\\User:TOR\\test	NaN
16348	\\Toxin_(Luminals)_(Earth-616)	Secret Identity
16349	\\Urizen_U1%27var_(Earth-616)	Public Identity
16350	\\Valka_(Earth-616)	NaN
16351	\\Vayor_(Earth-616)	No Dual Identity
16352	\\Viridian_(Earth-616)	NaN
16353	\\William_Burke_(Earth-616)	NaN
16354	\\William_Falsworth_(Earth-616)	NaN
16355	\\William_Shakespeare_(Earth-616)	Public Identity
16356	\\Yamata_no_Orichi_(Earth-616)	Public Identity
16357	\\Zero_G._Priestly_(Legion_Personality)_(Earth...	Secret Identity
16358	\\Zora_Loftus_(Earth-616)	Secret Identity
16359	\\Agar_(Earth-616)	No Dual Identity
16360	\\Ana_(Natasha_Romanoff%27s_neighbor)_(Earth-616)	No Dual Identity
16361	\\Dante%27s_mother_(Earth-616)	No Dual Identity
16362	\\Farbauti_(Earth-616)	Public Identity
16363	\\Finch_(Kate_Bishop%27s_neighbor)_(Earth-616)	Public Identity
16364	\\Jack_0%27Lantern_(Impostor)_(Earth-616)	Secret Identity
16365	\\K%27thol_(Earth-616)	NaN
16366	\\Karen_(Hijack%27s_girlfriend)_(Earth-616)	Public Identity
16367	\\Marcus_(Kate_Bishop%27s_neighbor)_(Earth-616)	Public Identity
16368	\\Marcy_(Offer%27s_employee)_(Earth-616)	Public Identity
16369	\\Melanie_Kapoor_(Earth-616)	Public Identity
16370	\\Phoenix%27s_Shadow_(Earth-616)	NaN
16371	\\Ru%27ach_(Earth-616)	No Dual Identity
16372	\\Thane_(Thanos%27_son)_(Earth-616)	No Dual Identity
16373	\\Tinkerer_(Skrull)_(Earth-616)	Secret Identity
16374	\\TK421_(Spiderling)_(Earth-616)	Secret Identity

16375

\Yologarch\_(Earth-616)

NaN

	ALIGN	EYE	HAIR	SEX \
0	Good Characters	Hazel Eyes	Brown Hair	Male Characters
1	Good Characters	Blue Eyes	White Hair	Male Characters
2	Neutral Characters	Blue Eyes	Black Hair	Male Characters
3	Good Characters	Blue Eyes	Black Hair	Male Characters
4	Good Characters	Blue Eyes	Blond Hair	Male Characters
5	Good Characters	Blue Eyes	No Hair	Male Characters
6	Good Characters	Brown Eyes	Brown Hair	Male Characters
7	Good Characters	Brown Eyes	Brown Hair	Male Characters
8	Neutral Characters	Brown Eyes	Brown Hair	Male Characters
9	Good Characters	Blue Eyes	Blond Hair	Male Characters
10	Good Characters	Blue Eyes	Blue Hair	Male Characters
11	Good Characters	Blue Eyes	Blond Hair	Female Characters
12	Neutral Characters	Green Eyes	Black Hair	Male Characters
13	Good Characters	Blue Eyes	White Hair	Female Characters
14	Good Characters	Blue Eyes	Blond Hair	Male Characters
15	Good Characters	Blue Eyes	Red Hair	Male Characters
16	Good Characters	Grey Eyes	Black Hair	Male Characters
17	Good Characters	Green Eyes	Red Hair	Female Characters
18	Neutral Characters	Blue Eyes	Black Hair	Male Characters
19	Good Characters	Brown Eyes	Brown Hair	Male Characters
20	Good Characters	Blue Eyes	Blond Hair	Male Characters
21	Good Characters	Blue Eyes	Bald	Male Characters
22	Good Characters	Blue Eyes	Blond Hair	Male Characters
23	Good Characters	Blue Eyes	Black Hair	Male Characters
24	Good Characters	Green Eyes	Brown Hair	Female Characters
25	Neutral Characters	Brown Eyes	Brown Hair	Male Characters
26	Good Characters	Blue Eyes	Auburn Hair	Female Characters
27	Good Characters	Green Eyes	Red Hair	Female Characters
28	Good Characters	Green Eyes	Red Hair	Female Characters
29	Good Characters	Yellow Eyes	Blue Hair	Male Characters
...	...	...	...	...
16346	NaN	NaN	NaN	Male Characters
16347	NaN	NaN	NaN	NaN
16348	NaN	NaN	No Hair	NaN
16349	Good Characters	White Eyes	NaN	Male Characters
16350	NaN	NaN	NaN	NaN
16351	Bad Characters	NaN	Black Hair	Male Characters
16352	NaN	NaN	NaN	Male Characters
16353	NaN	NaN	NaN	Male Characters
16354	Good Characters	NaN	NaN	Male Characters
16355	Neutral Characters	NaN	NaN	Male Characters
16356	Bad Characters	NaN	No Hair	Male Characters
16357	Neutral Characters	NaN	NaN	Male Characters
16358	Bad Characters	NaN	Blond Hair	Female Characters
16359	Bad Characters	Black Eyes	No Hair	Male Characters

16360	Good Characters	Black Eyes	Grey Hair	Female Characters
16361	Neutral Characters	NaN	NaN	Female Characters
16362	Bad Characters	Red Eyes	NaN	Female Characters
16363	Good Characters	Black Eyes	Bald	Male Characters
16364	Bad Characters	Hazel Eyes	Bald	Male Characters
16365	Good Characters	NaN	NaN	Male Characters
16366	Neutral Characters	Brown Eyes	Black Hair	Female Characters
16367	Good Characters	Hazel Eyes	Bald	Male Characters
16368	Neutral Characters	NaN	Brown Hair	Female Characters
16369	Good Characters	Blue Eyes	Black Hair	Female Characters
16370	Neutral Characters	NaN	NaN	NaN
16371	Bad Characters	Green Eyes	No Hair	Male Characters
16372	Good Characters	Blue Eyes	Bald	Male Characters
16373	Bad Characters	Black Eyes	Bald	Male Characters
16374	Neutral Characters	NaN	NaN	Male Characters
16375	Bad Characters	NaN	NaN	NaN

	GSM	ALIVE	APPEARANCES \
0	NaN	Living Characters	4043.0
1	NaN	Living Characters	3360.0
2	NaN	Living Characters	3061.0
3	NaN	Living Characters	2961.0
4	NaN	Living Characters	2258.0
5	NaN	Living Characters	2255.0
6	NaN	Living Characters	2072.0
7	NaN	Living Characters	2017.0
8	NaN	Living Characters	1955.0
9	NaN	Living Characters	1934.0
10	NaN	Living Characters	1825.0
11	NaN	Living Characters	1713.0
12	NaN	Living Characters	1528.0
13	NaN	Living Characters	1512.0
14	NaN	Living Characters	1394.0
15	NaN	Living Characters	1338.0
16	NaN	Living Characters	1307.0
17	NaN	Living Characters	1304.0
18	NaN	Living Characters	1266.0
19	NaN	Living Characters	1265.0
20	NaN	Living Characters	1237.0
21	NaN	Deceased Characters	1233.0
22	NaN	Living Characters	1230.0
23	NaN	Living Characters	1162.0
24	NaN	Living Characters	1161.0
25	NaN	Living Characters	1137.0
26	NaN	Living Characters	1120.0
27	NaN	Deceased Characters	1107.0
28	Bisexual Characters	Living Characters	1050.0
29	NaN	Living Characters	1047.0

...	...	...	...
16346	NaN	Deceased Characters	NaN
16347	NaN	NaN	NaN
16348	NaN	Living Characters	NaN
16349	NaN	Living Characters	NaN
16350	NaN	Living Characters	NaN
16351	NaN	Living Characters	NaN
16352	NaN	Living Characters	NaN
16353	NaN	Living Characters	NaN
16354	NaN	Deceased Characters	NaN
16355	NaN	Living Characters	NaN
16356	NaN	Living Characters	NaN
16357	NaN	Living Characters	NaN
16358	NaN	Living Characters	NaN
16359	NaN	Living Characters	NaN
16360	NaN	Living Characters	NaN
16361	NaN	Deceased Characters	NaN
16362	NaN	Living Characters	NaN
16363	Homosexual Characters	Living Characters	NaN
16364	NaN	Living Characters	NaN
16365	NaN	Deceased Characters	NaN
16366	NaN	Living Characters	NaN
16367	Homosexual Characters	Living Characters	NaN
16368	NaN	Living Characters	NaN
16369	NaN	Living Characters	NaN
16370	NaN	Living Characters	NaN
16371	NaN	Living Characters	NaN
16372	NaN	Living Characters	NaN
16373	NaN	Living Characters	NaN
16374	NaN	Living Characters	NaN
16375	NaN	Living Characters	NaN

	FIRST APPEARANCE	Year
0	Aug-62	1962.0
1	Mar-41	1941.0
2	Oct-74	1974.0
3	Mar-63	1963.0
4	Nov-50	1950.0
5	Nov-61	1961.0
6	Nov-61	1961.0
7	May-62	1962.0
8	Sep-63	1963.0
9	Nov-61	1961.0
10	Sep-63	1963.0
11	Nov-61	1961.0
12	NaN	NaN
13	May-75	1975.0
14	Sep-64	1964.0



15	Apr-64	1964.0
16	Jul-63	1963.0
17	Jun-65	1965.0
18	Mar-63	1963.0
19	Sep-63	1963.0
20	Jan-62	1962.0
21	Sep-63	1963.0
22	Sep-63	1963.0
23	May-75	1975.0
24	Mar-64	1964.0
25	May-63	1963.0
26	Jun-63	1963.0
27	Sep-63	1963.0
28	Apr-64	1964.0
29	May-75	1975.0
...	...	...
16346	NaN	NaN
16347	NaN	NaN
16348	NaN	NaN
16349	NaN	NaN
16350	NaN	NaN
16351	NaN	NaN
16352	NaN	NaN
16353	NaN	NaN
16354	NaN	NaN
16355	NaN	NaN
16356	NaN	NaN
16357	NaN	NaN
16358	NaN	NaN
16359	NaN	NaN
16360	NaN	NaN
16361	NaN	NaN
16362	NaN	NaN
16363	NaN	NaN
16364	NaN	NaN
16365	NaN	NaN
16366	NaN	NaN
16367	NaN	NaN
16368	NaN	NaN
16369	NaN	NaN
16370	NaN	NaN
16371	NaN	NaN
16372	NaN	NaN
16373	NaN	NaN
16374	NaN	NaN
16375	NaN	NaN

[16376 rows x 13 columns]

```
In [130]: data.isnull().sum()
```

```
Out[130]: page_id      0
          name         0
          urlslug      0
          ID           3770
          ALIGN        2812
          EYE          9767
          HAIR         4264
          SEX           854
          GSM          16286
          ALIVE         3
          APPEARANCES  1096
          FIRST APPEARANCE 815
          Year         815
          dtype: int64
```

```
In [7]: data.dtypes
```

```
Out[7]: page_id      int64
          name        object
          urlslug     object
          ID          object
          ALIGN       object
          EYE         object
          HAIR        object
          SEX         object
          GSM         object
          ALIVE       object
          APPEARANCES float64
          FIRST APPEARANCE object
          Year        float64
          dtype: object
```

1.1. -

```
In [8]: data_new_1 = data.dropna(axis=1, how='any')
          (data.shape, data_new_1.shape)
```

```
Out[8]: ((16376, 13), (16376, 3))
```

```
In [9]: data_new_2 = data.dropna(axis=0, how='any')
          (data.shape, data_new_2.shape)
```

```
Out[9]: ((16376, 13), (58, 13))
```

```
In [21]: data_new_3 = data.fillna(0)
          data_new_3.head(50)
```

```

Out[21]:
page_id      name \
0      1678      Spider-Man (Peter Parker)
1      7139      Captain America (Steven Rogers)
2      64786      Wolverine (James \"Logan\" Howlett)
3      1868      Iron Man (Anthony \"Tony\" Stark)
4      2460      Thor (Thor Odinson)
5      2458      Benjamin Grimm (Earth-616)
6      2166      Reed Richards (Earth-616)
7      1833      Hulk (Robert Bruce Banner)
8      29481      Scott Summers (Earth-616)
9      1837      Jonathan Storm (Earth-616)
10     15725      Henry McCoy (Earth-616)
11     1863      Susan Storm (Earth-616)
12     7823      Namor McKenzie (Earth-616)
13     2614      Ororo Munroe (Earth-616)
14     1803      Clinton Barton (Earth-616)
15     1396      Matthew Murdock (Earth-616)
16     55534      Stephen Strange (Earth-616)
17     1978      Mary Jane Watson (Earth-616)
18     1872      John Jonah Jameson (Earth-616)
19     35350      Robert Drake (Earth-616)
20     1557      Henry Pym (Earth-616)
21     65255      Charles Xavier (Earth-616)
22     1073      Warren Worthington III (Earth-616)
23     1346      Piotr Rasputin (Earth-616)
24     2512      Wanda Maximoff (Earth-616)
25     1671      Nicholas Fury (Earth-616)
26     1976      Janet van Dyne (Earth-616)
27     65219      Jean Grey (Earth-616)
28     6545      Natalia Romanova (Earth-616)
29     2223      Kurt Wagner (Earth-616)
30     2414      Vision (Earth-616)
31     8650      May Reilly (Earth-616)
32     2527      Katherine Pryde (Earth-616)
33     1970      Carol Danvers (Earth-616)
34     37690      Jennifer Walters (Earth-616)
35     3765      Emma Frost (Earth-616)
36     2312      Frank Castle (Earth-616)
37     1265      Luke Cage (Earth-616)
38     1677      Rogue (Anna Marie) (Earth-616)
39     25222      Conan (Earth-616)
40     6537      Joseph Robertson (Earth-616)
41     1592      Pietro Maximoff (Earth-616)
42     1818      Hercules (Earth-616)
43     1448      Victor von Doom (Earth-616)
44     2081      Max Eisenhardt (Earth-616)
45     2307      Elizabeth Braddock (Earth-616)
46     2556      Norrin Radd (Earth-616)

```

47	1740	Norman Osborn (Earth-616)
48	8651	Eugene Thompson (Earth-616)
49	2009	Simon Williams (Earth-616)

	urlslug	ID \
0	\\Spider-Man_(Peter_Parker)	Secret Identity
1	\\Captain_America_(Steven_Rogers)	Public Identity
2	\\Wolverine_(James_%22Logan%22_Howlett)	Public Identity
3	\\Iron_Man_(Anthony_%22Tony%22_Stark)	Public Identity
4	\\Thor_(Thor_Odinson)	No Dual Identity
5	\\Benjamin_Grimm_(Earth-616)	Public Identity
6	\\Reed_Richards_(Earth-616)	Public Identity
7	\\Hulk_(Robert_Bruce_Banner)	Public Identity
8	\\Scott_Summers_(Earth-616)	Public Identity
9	\\Jonathan_Storm_(Earth-616)	Public Identity
10	\\Henry_McCoy_(Earth-616)	Public Identity
11	\\Susan_Storm_(Earth-616)	Public Identity
12	\\Namor_McKenzie_(Earth-616)	No Dual Identity
13	\\Ororo_Munroe_(Earth-616)	Public Identity
14	\\Clinton_Barton_(Earth-616)	Public Identity
15	\\Matthew_Murdock_(Earth-616)	Public Identity
16	\\Stephen_Strange_(Earth-616)	Public Identity
17	\\Mary_Jane_Watson_(Earth-616)	No Dual Identity
18	\\John_Jonah_Jameson_(Earth-616)	No Dual Identity
19	\\Robert_Drake_(Earth-616)	Secret Identity
20	\\Henry_Pym_(Earth-616)	Public Identity
21	\\Charles_Xavier_(Earth-616)	Public Identity
22	\\Warren_Worthington_III_(Earth-616)	Public Identity
23	\\Piotr_Rasputin_(Earth-616)	Secret Identity
24	\\Wanda_Maximoff_(Earth-616)	Public Identity
25	\\Nicholas_Fury_(Earth-616)	No Dual Identity
26	\\Janet_van_Dyne_(Earth-616)	Public Identity
27	\\Jean_Grey_(Earth-616)	Public Identity
28	\\Natalia_Romanova_(Earth-616)	Public Identity
29	\\Kurt_Wagner_(Earth-616)	Secret Identity
30	\\Vision_(Earth-616)	Secret Identity
31	\\May_Reilly_(Earth-616)	No Dual Identity
32	\\Katherine_Pryde_(Earth-616)	Secret Identity
33	\\Carol_Danvers_(Earth-616)	Public Identity
34	\\Jennifer_Walters_(Earth-616)	Public Identity
35	\\Emma_Frost_(Earth-616)	Public Identity
36	\\Frank_Castle_(Earth-616)	Public Identity
37	\\Luke_Cage_(Earth-616)	No Dual Identity
38	\\Rogue_(Anna_Marie)_(Earth-616)	Secret Identity
39	\\Conan_(Earth-616)	No Dual Identity
40	\\Joseph_Robertson_(Earth-616)	No Dual Identity
41	\\Pietro_Maximoff_(Earth-616)	Public Identity
42	\\Hercules_(Earth-616)	No Dual Identity

43	\\Victor_von_Doom_(Earth-616)	Public Identity
44	\\Max_Eisenhardt_(Earth-616)	Public Identity
45	\\Elizabeth_Braddock_(Earth-616)	Secret Identity
46	\\Norrrin_Radd_(Earth-616)	Public Identity
47	\\Norman_Osborn_(Earth-616)	Public Identity
48	\\Eugene_Thompson_(Earth-616)	Secret Identity
49	\\Simon_Williams_(Earth-616)	Public Identity

	ALIGN	EYE	HAIR	SEX	\
0	Good Characters	Hazel Eyes	Brown Hair	Male Characters	
1	Good Characters	Blue Eyes	White Hair	Male Characters	
2	Neutral Characters	Blue Eyes	Black Hair	Male Characters	
3	Good Characters	Blue Eyes	Black Hair	Male Characters	
4	Good Characters	Blue Eyes	Blond Hair	Male Characters	
5	Good Characters	Blue Eyes	No Hair	Male Characters	
6	Good Characters	Brown Eyes	Brown Hair	Male Characters	
7	Good Characters	Brown Eyes	Brown Hair	Male Characters	
8	Neutral Characters	Brown Eyes	Brown Hair	Male Characters	
9	Good Characters	Blue Eyes	Blond Hair	Male Characters	
10	Good Characters	Blue Eyes	Blue Hair	Male Characters	
11	Good Characters	Blue Eyes	Blond Hair	Female Characters	
12	Neutral Characters	Green Eyes	Black Hair	Male Characters	
13	Good Characters	Blue Eyes	White Hair	Female Characters	
14	Good Characters	Blue Eyes	Blond Hair	Male Characters	
15	Good Characters	Blue Eyes	Red Hair	Male Characters	
16	Good Characters	Grey Eyes	Black Hair	Male Characters	
17	Good Characters	Green Eyes	Red Hair	Female Characters	
18	Neutral Characters	Blue Eyes	Black Hair	Male Characters	
19	Good Characters	Brown Eyes	Brown Hair	Male Characters	
20	Good Characters	Blue Eyes	Blond Hair	Male Characters	
21	Good Characters	Blue Eyes	Bald	Male Characters	
22	Good Characters	Blue Eyes	Blond Hair	Male Characters	
23	Good Characters	Blue Eyes	Black Hair	Male Characters	
24	Good Characters	Green Eyes	Brown Hair	Female Characters	
25	Neutral Characters	Brown Eyes	Brown Hair	Male Characters	
26	Good Characters	Blue Eyes	Auburn Hair	Female Characters	
27	Good Characters	Green Eyes	Red Hair	Female Characters	
28	Good Characters	Green Eyes	Red Hair	Female Characters	
29	Good Characters	Yellow Eyes	Blue Hair	Male Characters	
30	Good Characters	Gold Eyes	No Hair	Male Characters	
31	Good Characters	Blue Eyes	Grey Hair	Female Characters	
32	Good Characters	Hazel Eyes	Brown Hair	Female Characters	
33	Good Characters	Blue Eyes	Blond Hair	Female Characters	
34	Good Characters	Green Eyes	Brown Hair	Female Characters	
35	Neutral Characters	Blue Eyes	Brown Hair	Female Characters	
36	Neutral Characters	Blue Eyes	Black Hair	Male Characters	
37	Good Characters	Brown Eyes	Black Hair	Male Characters	
38	Good Characters	Green Eyes	Auburn Hair	Female Characters	

39	Neutral Characters	Blue Eyes	Black Hair	Male Characters
40	Good Characters	Brown Eyes	White Hair	Male Characters
41	Good Characters	Blue Eyes	Silver Hair	Male Characters
42	Good Characters	Blue Eyes	Brown Hair	Male Characters
43	Bad Characters	Brown Eyes	Brown Hair	Male Characters
44	Neutral Characters	Grey Eyes	Bald	Male Characters
45	Neutral Characters	Blue Eyes	Purple Hair	Female Characters
46	Good Characters	Blue Eyes	Bald	Male Characters
47	Bad Characters	Blue Eyes	Auburn Hair	Male Characters
48	Good Characters	Blue Eyes	Strawberry Blond Hair	Male Characters
49	Good Characters	Red Eyes	Grey Hair	Male Characters

	GSM	ALIVE	APPEARANCES	FIRST APPEARANCE	\
0	0	Living Characters	4043.0	Aug-62	
1	0	Living Characters	3360.0	Mar-41	
2	0	Living Characters	3061.0	Oct-74	
3	0	Living Characters	2961.0	Mar-63	
4	0	Living Characters	2258.0	Nov-50	
5	0	Living Characters	2255.0	Nov-61	
6	0	Living Characters	2072.0	Nov-61	
7	0	Living Characters	2017.0	May-62	
8	0	Living Characters	1955.0	Sep-63	
9	0	Living Characters	1934.0	Nov-61	
10	0	Living Characters	1825.0	Sep-63	
11	0	Living Characters	1713.0	Nov-61	
12	0	Living Characters	1528.0		0
13	0	Living Characters	1512.0	May-75	
14	0	Living Characters	1394.0	Sep-64	
15	0	Living Characters	1338.0	Apr-64	
16	0	Living Characters	1307.0	Jul-63	
17	0	Living Characters	1304.0	Jun-65	
18	0	Living Characters	1266.0	Mar-63	
19	0	Living Characters	1265.0	Sep-63	
20	0	Living Characters	1237.0	Jan-62	
21	0	Deceased Characters	1233.0	Sep-63	
22	0	Living Characters	1230.0	Sep-63	
23	0	Living Characters	1162.0	May-75	
24	0	Living Characters	1161.0	Mar-64	
25	0	Living Characters	1137.0	May-63	
26	0	Living Characters	1120.0	Jun-63	
27	0	Deceased Characters	1107.0	Sep-63	
28	Bisexual Characters	Living Characters	1050.0	Apr-64	
29		Living Characters	1047.0	May-75	
30		Living Characters	1007.0	Oct-68	
31		Living Characters	1000.0	Aug-62	
32		Living Characters	886.0	Jan-80	
33		Living Characters	881.0	Mar-68	
34		Living Characters	881.0	Feb-80	

35	0	Living Characters	880.0	Jan-80
36	0	Living Characters	878.0	Feb-74
37	0	Living Characters	856.0	Jun-72
38	0	Living Characters	850.0	0
39	0	Living Characters	787.0	Oct-70
40	0	Living Characters	752.0	Aug-67
41	0	Living Characters	740.0	Mar-64
42	Bisexual Characters	Living Characters	725.0	Oct-65
43	0	Living Characters	721.0	Jul-62
44	0	Living Characters	709.0	Sep-63
45	Bisexual Characters	Living Characters	708.0	Dec-76
46	0	Living Characters	696.0	Mar-66
47	0	Living Characters	692.0	Jul-64
48	0	Living Characters	689.0	Aug-62
49	0	Living Characters	686.0	Oct-64

	Year
0	1962.0
1	1941.0
2	1974.0
3	1963.0
4	1950.0
5	1961.0
6	1961.0
7	1962.0
8	1963.0
9	1961.0
10	1963.0
11	1961.0
12	0.0
13	1975.0
14	1964.0
15	1964.0
16	1963.0
17	1965.0
18	1963.0
19	1963.0
20	1962.0
21	1963.0
22	1963.0
23	1975.0
24	1964.0
25	1963.0
26	1963.0
27	1963.0
28	1964.0
29	1975.0
30	1968.0

```

31 1962.0
32 1980.0
33 1968.0
34 1980.0
35 1980.0
36 1974.0
37 1972.0
38 0.0
39 1970.0
40 1967.0
41 1964.0
42 1965.0
43 1962.0
44 1963.0
45 1976.0
46 1966.0
47 1964.0
48 1962.0
49 1964.0

```

1.2. " " - (imputation) 1.2.1.

```

In [131]: #
#
num_cols = []
for col in data.columns:
    #
    temp_null_count = data[data[col].isnull()].shape[0]
    dt = str(data[col].dtype)
    if temp_null_count>0 and (dt=='float64' or dt=='int64'):
        num_cols.append(col)
        print(' {}.  {}.  {}'.format(col, dt, temp_null_count))

```

```

APPEARANCES.  float64.  1096.
Year.  float64.  815.

```

```

In [132]: data_num = data[num_cols]
data_num

```

```

Out[132]:

```

	APPEARANCES	Year
0	4043.0	1962.0
1	3360.0	1941.0
2	3061.0	1974.0
3	2961.0	1963.0
4	2258.0	1950.0
5	2255.0	1961.0
6	2072.0	1961.0
7	2017.0	1962.0



8	1955.0	1963.0
9	1934.0	1961.0
10	1825.0	1963.0
11	1713.0	1961.0
12	1528.0	NaN
13	1512.0	1975.0
14	1394.0	1964.0
15	1338.0	1964.0
16	1307.0	1963.0
17	1304.0	1965.0
18	1266.0	1963.0
19	1265.0	1963.0
20	1237.0	1962.0
21	1233.0	1963.0
22	1230.0	1963.0
23	1162.0	1975.0
24	1161.0	1964.0
25	1137.0	1963.0
26	1120.0	1963.0
27	1107.0	1963.0
28	1050.0	1964.0
29	1047.0	1975.0
...	...	...
16346	NaN	NaN
16347	NaN	NaN
16348	NaN	NaN
16349	NaN	NaN
16350	NaN	NaN
16351	NaN	NaN
16352	NaN	NaN
16353	NaN	NaN
16354	NaN	NaN
16355	NaN	NaN
16356	NaN	NaN
16357	NaN	NaN
16358	NaN	NaN
16359	NaN	NaN
16360	NaN	NaN
16361	NaN	NaN
16362	NaN	NaN
16363	NaN	NaN
16364	NaN	NaN
16365	NaN	NaN
16366	NaN	NaN
16367	NaN	NaN
16368	NaN	NaN
16369	NaN	NaN
16370	NaN	NaN

16371	NaN	NaN
16372	NaN	NaN
16373	NaN	NaN
16374	NaN	NaN
16375	NaN	NaN

[16376 rows x 2 columns]

```
In [68]: # MasVnrArea
data[data['Year'].isnull()]
```

```
Out[68]:
```

	page_id	name \
12	7823	Namor McKenzie (Earth-616)
38	1677	Rogue (Anna Marie) (Earth-616)
80	67048	Blaine Colt (Earth-616)
114	37751	Monica Rambeau (Earth-616)
259	25255	James Bradley (Earth-616)
310	535259	Steel (Earth-616)
413	626857	Howard Hanover (Earth-616)
683	13066	Fen (Earth-616)
789	103136	Thakorr (Earth-616)
854	30883	Brightwind (Earth-616)
997	18445	Ahura Boltagon (Earth-616)
1118	17489	Francesca Grace (Earth-616)
1158	1771	Alistaire Smythe (Earth-616)
1316	2126	Martinique Wyngarde (Earth-616)
1454	93233	Leonard McKenzie (Earth-616)
1564	40064	Hrimhari (Earth-616)
1565	638526	David Bond (Earth-616)
1845	631202	Aldrif Odinsdottir (Earth-616)
1937	655072	Shogo Lee (Earth-616)
2033	6632	Leila Davis (Earth-616)
2147	665003	Theodore Marley Brooks (Earth-616)
2148	670564	Arno Stark (Earth-616)
2398	694425	John Renwick (Earth-616)
2399	640754	Proxima Midnight (Earth-616)
2549	21721	Perrikus (Earth-616)
2727	635523	Doombot (Avenger) (Earth-616)
2728	678469	Selah Burke (Earth-616)
2909	17957	Teresa Martin (Earth-616)
2910	635275	Alexis (Earth-616)
2911	674895	Kamala Khan (Earth-616)
...	...	...
16346	42106	Thiazzi (Earth-616)
16347	462671	TOR\test
16348	725312	Toxin (Luminals) (Earth-616)
16349	507239	Urizen Ul'var (Earth-616)
16350	40061	Valka (Earth-616)

16351	505759	Vayor (Earth-616)
16352	16569	Viridian (Earth-616)
16353	16296	William Burke (Earth-616)
16354	120833	William Falsworth (Earth-616)
16355	693077	William Shakespeare (Earth-616)
16356	473710	Yamata no Orichi (Earth-616)
16357	670512	Zero G. Priestly (Legion Personality) (Earth-616)
16358	12132	Zora Loftus (Earth-616)
16359	727035	Agar (Earth-616)
16360	703546	Ana (Natasha Romanoff's neighbor) (Earth-616)
16361	713903	Dante's mother (Earth-616)
16362	508693	Farbauti (Earth-616)
16363	694577	Finch (Kate Bishop's neighbor) (Earth-616)
16364	655292	Jack O'Lantern (Impostor) (Earth-616)
16365	684262	K'thol (Earth-616)
16366	643435	Karen (Hijack's girlfriend) (Earth-616)
16367	694583	Marcus (Kate Bishop's neighbor) (Earth-616)
16368	703892	Marcy (Offer's employee) (Earth-616)
16369	660799	Melanie Kapoor (Earth-616)
16370	674414	Phoenix's Shadow (Earth-616)
16371	657508	Ru'ach (Earth-616)
16372	665474	Thane (Thanos' son) (Earth-616)
16373	695217	Tinkerer (Skrull) (Earth-616)
16374	708811	TK421 (Spiderling) (Earth-616)
16375	673702	Yologarch (Earth-616)

	urlslug	ID \
12	\\Namor_McKenzie_(Earth-616)	No Dual Identity
38	\\Rogue_(Anna_Marie)_(Earth-616)	Secret Identity
80	\\Blaine_Colt_(Earth-616)	Public Identity
114	\\Monica_Rambeau_(Earth-616)	Secret Identity
259	\\James_Bradley_(Earth-616)	Secret Identity
310	\\Steel_(Earth-616)	Public Identity
413	\\Howard_Hanover_(Earth-616)	Public Identity
683	\\Fen_(Earth-616)	No Dual Identity
789	\\Thakorr_(Earth-616)	No Dual Identity
854	\\Brightwind_(Earth-616)	NaN
997	\\Ahura_Boltagon_(Earth-616)	No Dual Identity
1118	\\Francesca_Grace_(Earth-616)	Secret Identity
1158	\\Alistaire_Smythe_(Earth-616)	Public Identity
1316	\\Martinique_Wyngarde_(Earth-616)	Secret Identity
1454	\\Leonard_McKenzie_(Earth-616)	No Dual Identity
1564	\\Hrimhari_(Earth-616)	Secret Identity
1565	\\David_Bond_(Earth-616)	Secret Identity
1845	\\Aldrif_Odinsdottir_(Earth-616)	Secret Identity
1937	\\Shogo_Lee_(Earth-616)	No Dual Identity
2033	\\Leila_Davis_(Earth-616)	Public Identity
2147	\\Theodore_Marley_Brooks_(Earth-616)	No Dual Identity

2148	\\Arno_Stark_(Earth-616)	No Dual Identity
2398	\\John_Renwick_(Earth-616)	No Dual Identity
2399	\\Proxima_Midnight_(Earth-616)	No Dual Identity
2549	\\Perrikus_(Earth-616)	NaN
2727	\\Doombot_(Avenger)_(Earth-616)	Public Identity
2728	\\Selah_Burke_(Earth-616)	Secret Identity
2909	\\Teresa_Martin_(Earth-616)	Public Identity
2910	\\Alexis_(Earth-616)	Public Identity
2911	\\Kamala_Khan_(Earth-616)	Secret Identity
...	...	...
16346	\\Thiazzi_(Earth-616)	NaN
16347	\\User:TOR\\test	NaN
16348	\\Toxin_(Luminals)_(Earth-616)	Secret Identity
16349	\\Urizen_UL%27var_(Earth-616)	Public Identity
16350	\\Valka_(Earth-616)	NaN
16351	\\Vayor_(Earth-616)	No Dual Identity
16352	\\Viridian_(Earth-616)	NaN
16353	\\William_Burke_(Earth-616)	NaN
16354	\\William_Falsworth_(Earth-616)	NaN
16355	\\William_Shakespeare_(Earth-616)	Public Identity
16356	\\Yamata_no_Orichi_(Earth-616)	Public Identity
16357	\\Zero_G._Priestly_(Legion_Personality)_(Earth...	Secret Identity
16358	\\Zora_Loftus_(Earth-616)	Secret Identity
16359	\\Agar_(Earth-616)	No Dual Identity
16360	\\Ana_(Natasha_Romanoff%27s_neighbor)_(Earth-616)	No Dual Identity
16361	\\Dante%27s_mother_(Earth-616)	No Dual Identity
16362	\\Farbauti_(Earth-616)	Public Identity
16363	\\Finch_(Kate_Bishop%27s_neighbor)_(Earth-616)	Public Identity
16364	\\Jack_0%27Lantern_(Impostor)_(Earth-616)	Secret Identity
16365	\\K%27thol_(Earth-616)	NaN
16366	\\Karen_(Hijack%27s_girlfriend)_(Earth-616)	Public Identity
16367	\\Marcus_(Kate_Bishop%27s_neighbor)_(Earth-616)	Public Identity
16368	\\Marcy_(Offer%27s_employee)_(Earth-616)	Public Identity
16369	\\Melanie_Kapoor_(Earth-616)	Public Identity
16370	\\Phoenix%27s_Shadow_(Earth-616)	NaN
16371	\\Ru%27ach_(Earth-616)	No Dual Identity
16372	\\Thane_(Thanos%27_son)_(Earth-616)	No Dual Identity
16373	\\Tinkerer_(Skrull)_(Earth-616)	Secret Identity
16374	\\TK421_(Spiderling)_(Earth-616)	Secret Identity
16375	\\Yologarch_(Earth-616)	NaN

	ALIGN	EYE	HAIR	SEX \
12	Neutral Characters	Green Eyes	Black Hair	Male Characters
38	Good Characters	Green Eyes	Auburn Hair	Female Characters
80	NaN	Blue Eyes	Blond Hair	Male Characters
114	Good Characters	Brown Eyes	Black Hair	Female Characters
259	Good Characters	Blue Eyes	Black Hair	Male Characters
310	Good Characters	NaN	White Hair	Male Characters

413		NaN	NaN	Black Hair	Male Characters
683	Good Characters	Green Eyes	Auburn Hair	Female Characters	
789	Bad Characters	NaN	No Hair	Male Characters	
854		NaN	Red Eyes	Black Hair	Male Characters
997		NaN	Brown Eyes	Black Hair	Male Characters
1118		NaN	Black Eyes	Black Hair	Female Characters
1158	Bad Characters	Brown Eyes	Brown Hair	Male Characters	
1316	Bad Characters	Blue Eyes	Brown Hair	Female Characters	
1454	Good Characters	Brown Eyes	Black Hair	Male Characters	
1564	Good Characters	Silver Eyes	Grey Hair	Male Characters	
1565	Good Characters	Brown Eyes	Black Hair	Male Characters	
1845	Neutral Characters	White Eyes	Red Hair	Female Characters	
1937	Neutral Characters	Brown Eyes	Black Hair	Male Characters	
2033	Neutral Characters	Blue Eyes	Black Hair	Female Characters	
2147	Good Characters	NaN	NaN	Male Characters	
2148	Good Characters	Blue Eyes	Black Hair	Male Characters	
2398	Good Characters	NaN	NaN	Male Characters	
2399	Bad Characters	White Eyes	Blue Hair	Female Characters	
2549	Bad Characters	Red Eyes	NaN	Male Characters	
2727	Good Characters	Red Eyes	No Hair	Male Characters	
2728	Good Characters	Brown Eyes	Black Hair	Female Characters	
2909	Good Characters	Brown Eyes	Brown Hair	Female Characters	
2910	Good Characters	Purple Eyes	Blond Hair	Female Characters	
2911	Good Characters	Brown Eyes	Brown Hair	Female Characters	
...		...	...	...	...
16346		NaN	NaN	NaN	Male Characters
16347		NaN	NaN	NaN	NaN
16348		NaN	NaN	No Hair	NaN
16349	Good Characters	White Eyes	NaN	Male Characters	
16350		NaN	NaN	NaN	NaN
16351	Bad Characters	NaN	Black Hair	Male Characters	
16352		NaN	NaN	Male Characters	
16353		NaN	NaN	Male Characters	
16354	Good Characters	NaN	NaN	Male Characters	
16355	Neutral Characters	NaN	NaN	Male Characters	
16356	Bad Characters	NaN	No Hair	Male Characters	
16357	Neutral Characters	NaN	NaN	Male Characters	
16358	Bad Characters	NaN	Blond Hair	Female Characters	
16359	Bad Characters	Black Eyes	No Hair	Male Characters	
16360	Good Characters	Black Eyes	Grey Hair	Female Characters	
16361	Neutral Characters	NaN	NaN	Female Characters	
16362	Bad Characters	Red Eyes	NaN	Female Characters	
16363	Good Characters	Black Eyes	Bald	Male Characters	
16364	Bad Characters	Hazel Eyes	Bald	Male Characters	
16365	Good Characters	NaN	NaN	Male Characters	
16366	Neutral Characters	Brown Eyes	Black Hair	Female Characters	
16367	Good Characters	Hazel Eyes	Bald	Male Characters	
16368	Neutral Characters	NaN	Brown Hair	Female Characters	

16369	Good Characters	Blue Eyes	Black Hair	Female Characters
16370	Neutral Characters	NaN	NaN	NaN
16371	Bad Characters	Green Eyes	No Hair	Male Characters
16372	Good Characters	Blue Eyes	Bald	Male Characters
16373	Bad Characters	Black Eyes	Bald	Male Characters
16374	Neutral Characters	NaN	NaN	Male Characters
16375	Bad Characters	NaN	NaN	NaN

	GSM	ALIVE	APPEARANCES \
12	NaN	Living Characters	1528.0
38	NaN	Living Characters	850.0
80	NaN	Deceased Characters	429.0
114	NaN	Living Characters	327.0
259	NaN	Living Characters	158.0
310	NaN	Living Characters	134.0
413	NaN	Living Characters	100.0
683	NaN	Living Characters	55.0
789	NaN	Deceased Characters	48.0
854	NaN	Living Characters	44.0
997	NaN	Living Characters	38.0
1118	NaN	Living Characters	34.0
1158	NaN	Deceased Characters	33.0
1316	NaN	Living Characters	29.0
1454	NaN	Deceased Characters	26.0
1564	NaN	Deceased Characters	24.0
1565	NaN	Living Characters	24.0
1845	NaN	Living Characters	20.0
1937	NaN	Living Characters	19.0
2033	NaN	Deceased Characters	18.0
2147	NaN	Living Characters	17.0
2148	NaN	Living Characters	17.0
2398	NaN	Living Characters	15.0
2399	NaN	Living Characters	15.0
2549	NaN	Living Characters	14.0
2727	NaN	Living Characters	13.0
2728	NaN	Living Characters	13.0
2909	NaN	Living Characters	12.0
2910	NaN	Living Characters	12.0
2911	NaN	Living Characters	12.0
...	...	...	...
16346	NaN	Deceased Characters	NaN
16347	NaN	NaN	NaN
16348	NaN	Living Characters	NaN
16349	NaN	Living Characters	NaN
16350	NaN	Living Characters	NaN
16351	NaN	Living Characters	NaN
16352	NaN	Living Characters	NaN
16353	NaN	Living Characters	NaN

16354		NaN	Deceased Characters	NaN
16355		NaN	Living Characters	NaN
16356		NaN	Living Characters	NaN
16357		NaN	Living Characters	NaN
16358		NaN	Living Characters	NaN
16359		NaN	Living Characters	NaN
16360		NaN	Living Characters	NaN
16361		NaN	Deceased Characters	NaN
16362		NaN	Living Characters	NaN
16363	Homosexual Characters		Living Characters	NaN
16364		NaN	Living Characters	NaN
16365		NaN	Deceased Characters	NaN
16366		NaN	Living Characters	NaN
16367	Homosexual Characters		Living Characters	NaN
16368		NaN	Living Characters	NaN
16369		NaN	Living Characters	NaN
16370		NaN	Living Characters	NaN
16371		NaN	Living Characters	NaN
16372		NaN	Living Characters	NaN
16373		NaN	Living Characters	NaN
16374		NaN	Living Characters	NaN
16375		NaN	Living Characters	NaN

	FIRST APPEARANCE	Year
12	NaN	NaN
38	NaN	NaN
80	NaN	NaN
114	NaN	NaN
259	NaN	NaN
310	NaN	NaN
413	NaN	NaN
683	NaN	NaN
789	NaN	NaN
854	NaN	NaN
997	NaN	NaN
1118	NaN	NaN
1158	NaN	NaN
1316	NaN	NaN
1454	NaN	NaN
1564	NaN	NaN
1565	NaN	NaN
1845	NaN	NaN
1937	NaN	NaN
2033	NaN	NaN
2147	NaN	NaN
2148	NaN	NaN
2398	NaN	NaN
2399	NaN	NaN

2549	NaN	NaN
2727	NaN	NaN
2728	NaN	NaN
2909	NaN	NaN
2910	NaN	NaN
2911	NaN	NaN
...	...	...
16346	NaN	NaN
16347	NaN	NaN
16348	NaN	NaN
16349	NaN	NaN
16350	NaN	NaN
16351	NaN	NaN
16352	NaN	NaN
16353	NaN	NaN
16354	NaN	NaN
16355	NaN	NaN
16356	NaN	NaN
16357	NaN	NaN
16358	NaN	NaN
16359	NaN	NaN
16360	NaN	NaN
16361	NaN	NaN
16362	NaN	NaN
16363	NaN	NaN
16364	NaN	NaN
16365	NaN	NaN
16366	NaN	NaN
16367	NaN	NaN
16368	NaN	NaN
16369	NaN	NaN
16370	NaN	NaN
16371	NaN	NaN
16372	NaN	NaN
16373	NaN	NaN
16374	NaN	NaN
16375	NaN	NaN

[815 rows x 13 columns]

In [75]: #

```
flt_index = data[data['Year'].isnull()].index
flt_index
```

Out[75]: Int64Index([ 12, 38, 80, 114, 259, 310, 413, 683, 789, 854, ..., 16366, 16367, 16368, 16369, 16370, 16371, 16372, 16373, 16374,



```
16375],
dtype='int64', length=815)
```

```
In [76]: #
         data[data.index.isin(flt_index)]
```

```
Out [76]:
```

	page_id	name \
12	7823	Namor McKenzie (Earth-616)
38	1677	Rogue (Anna Marie) (Earth-616)
80	67048	Blaine Colt (Earth-616)
114	37751	Monica Rambeau (Earth-616)
259	25255	James Bradley (Earth-616)
310	535259	Steel (Earth-616)
413	626857	Howard Hanover (Earth-616)
683	13066	Fen (Earth-616)
789	103136	Thakorr (Earth-616)
854	30883	Brightwind (Earth-616)
997	18445	Ahura Boltagon (Earth-616)
1118	17489	Francesca Grace (Earth-616)
1158	1771	Alistaire Smythe (Earth-616)
1316	2126	Martinique Wyngarde (Earth-616)
1454	93233	Leonard McKenzie (Earth-616)
1564	40064	Hrimhari (Earth-616)
1565	638526	David Bond (Earth-616)
1845	631202	Aldrif Odinsdottir (Earth-616)
1937	655072	Shogo Lee (Earth-616)
2033	6632	Leila Davis (Earth-616)
2147	665003	Theodore Marley Brooks (Earth-616)
2148	670564	Arno Stark (Earth-616)
2398	694425	John Renwick (Earth-616)
2399	640754	Proxima Midnight (Earth-616)
2549	21721	Perrikus (Earth-616)
2727	635523	Doombot (Avenger) (Earth-616)
2728	678469	Selah Burke (Earth-616)
2909	17957	Teresa Martin (Earth-616)
2910	635275	Alexis (Earth-616)
2911	674895	Kamala Khan (Earth-616)
...	...	...
16346	42106	Thiazzi (Earth-616)
16347	462671	TOR\test
16348	725312	Toxin (Luminals) (Earth-616)
16349	507239	Urizen Ul'var (Earth-616)
16350	40061	Valka (Earth-616)
16351	505759	Vayor (Earth-616)
16352	16569	Viridian (Earth-616)
16353	16296	William Burke (Earth-616)
16354	120833	William Falsworth (Earth-616)
16355	693077	William Shakespeare (Earth-616)

16356	473710	Yamata no Orichi (Earth-616)
16357	670512	Zero G. Priestly (Legion Personality) (Earth-616)
16358	12132	Zora Loftus (Earth-616)
16359	727035	Agar (Earth-616)
16360	703546	Ana (Natasha Romanoff's neighbor) (Earth-616)
16361	713903	Dante's mother (Earth-616)
16362	508693	Farbauti (Earth-616)
16363	694577	Finch (Kate Bishop's neighbor) (Earth-616)
16364	655292	Jack O'Lantern (Impostor) (Earth-616)
16365	684262	K'thol (Earth-616)
16366	643435	Karen (Hijack's girlfriend) (Earth-616)
16367	694583	Marcus (Kate Bishop's neighbor) (Earth-616)
16368	703892	Marcy (Offer's employee) (Earth-616)
16369	660799	Melanie Kapoor (Earth-616)
16370	674414	Phoenix's Shadow (Earth-616)
16371	657508	Ru'ach (Earth-616)
16372	665474	Thane (Thanos' son) (Earth-616)
16373	695217	Tinkerer (Skrull) (Earth-616)
16374	708811	TK421 (Spiderling) (Earth-616)
16375	673702	Yologarch (Earth-616)

	urlslug	ID \
12	\\Namor_McKenzie_(Earth-616)	No Dual Identity
38	\\Rogue_(Anna_Marie)_(Earth-616)	Secret Identity
80	\\Blaine_Colt_(Earth-616)	Public Identity
114	\\Monica_Rambeau_(Earth-616)	Secret Identity
259	\\James_Bradley_(Earth-616)	Secret Identity
310	\\Steel_(Earth-616)	Public Identity
413	\\Howard_Hanover_(Earth-616)	Public Identity
683	\\Fen_(Earth-616)	No Dual Identity
789	\\Thakorr_(Earth-616)	No Dual Identity
854	\\Brightwind_(Earth-616)	NaN
997	\\Ahura_Boltagon_(Earth-616)	No Dual Identity
1118	\\Francesca_Grace_(Earth-616)	Secret Identity
1158	\\Alistaire_Smythe_(Earth-616)	Public Identity
1316	\\Martinique_Wyngarde_(Earth-616)	Secret Identity
1454	\\Leonard_McKenzie_(Earth-616)	No Dual Identity
1564	\\Hrimhari_(Earth-616)	Secret Identity
1565	\\David_Bond_(Earth-616)	Secret Identity
1845	\\Aldrif_Odinsdottir_(Earth-616)	Secret Identity
1937	\\Shogo_Lee_(Earth-616)	No Dual Identity
2033	\\Leila_Davis_(Earth-616)	Public Identity
2147	\\Theodore_Marley_Brooks_(Earth-616)	No Dual Identity
2148	\\Arno_Stark_(Earth-616)	No Dual Identity
2398	\\John_Renwick_(Earth-616)	No Dual Identity
2399	\\Proxima_Midnight_(Earth-616)	No Dual Identity
2549	\\Perrikus_(Earth-616)	NaN
2727	\\Doombot_(Avenger)_(Earth-616)	Public Identity

2728	\\Selah_Burke_(Earth-616)	Secret Identity
2909	\\Teresa_Martin_(Earth-616)	Public Identity
2910	\\Alexis_(Earth-616)	Public Identity
2911	\\Kamala_Khan_(Earth-616)	Secret Identity
...	...	...
16346	\\Thiazzi_(Earth-616)	NaN
16347	\\User:TOR\\test	NaN
16348	\\Toxin_(Luminals)_(Earth-616)	Secret Identity
16349	\\Urizen_Ul%27var_(Earth-616)	Public Identity
16350	\\Valka_(Earth-616)	NaN
16351	\\Vayor_(Earth-616)	No Dual Identity
16352	\\Viridian_(Earth-616)	NaN
16353	\\William_Burke_(Earth-616)	NaN
16354	\\William_Falsworth_(Earth-616)	NaN
16355	\\William_Shakespeare_(Earth-616)	Public Identity
16356	\\Yamata_no_Orichi_(Earth-616)	Public Identity
16357	\\Zero_G._Priestly_(Legion_Personality)_(Earth...	Secret Identity
16358	\\Zora_Loftus_(Earth-616)	Secret Identity
16359	\\Agar_(Earth-616)	No Dual Identity
16360	\\Ana_(Natasha_Romanoff%27s_neighbor)_(Earth-616)	No Dual Identity
16361	\\Dante%27s_mother_(Earth-616)	No Dual Identity
16362	\\Farbauti_(Earth-616)	Public Identity
16363	\\Finch_(Kate_Bishop%27s_neighbor)_(Earth-616)	Public Identity
16364	\\Jack_0%27Lantern_(Impostor)_(Earth-616)	Secret Identity
16365	\\K%27thol_(Earth-616)	NaN
16366	\\Karen_(Hijack%27s_girlfriend)_(Earth-616)	Public Identity
16367	\\Marcus_(Kate_Bishop%27s_neighbor)_(Earth-616)	Public Identity
16368	\\Marcy_(Offer%27s_employee)_(Earth-616)	Public Identity
16369	\\Melanie_Kapoor_(Earth-616)	Public Identity
16370	\\Phoenix%27s_Shadow_(Earth-616)	NaN
16371	\\Ru%27ach_(Earth-616)	No Dual Identity
16372	\\Thane_(Thanos%27_son)_(Earth-616)	No Dual Identity
16373	\\Tinkerer_(Skrull)_(Earth-616)	Secret Identity
16374	\\TK421_(Spiderling)_(Earth-616)	Secret Identity
16375	\\Yologarch_(Earth-616)	NaN

	ALIGN	EYE	HAIR	SEX \
12	Neutral Characters	Green Eyes	Black Hair	Male Characters
38	Good Characters	Green Eyes	Auburn Hair	Female Characters
80	NaN	Blue Eyes	Blond Hair	Male Characters
114	Good Characters	Brown Eyes	Black Hair	Female Characters
259	Good Characters	Blue Eyes	Black Hair	Male Characters
310	Good Characters	NaN	White Hair	Male Characters
413	NaN	NaN	Black Hair	Male Characters
683	Good Characters	Green Eyes	Auburn Hair	Female Characters
789	Bad Characters	NaN	No Hair	Male Characters
854	NaN	Red Eyes	Black Hair	Male Characters
997	NaN	Brown Eyes	Black Hair	Male Characters

1118		NaN	Black Eyes	Black Hair	Female Characters
1158	Bad Characters		Brown Eyes	Brown Hair	Male Characters
1316	Bad Characters		Blue Eyes	Brown Hair	Female Characters
1454	Good Characters		Brown Eyes	Black Hair	Male Characters
1564	Good Characters		Silver Eyes	Grey Hair	Male Characters
1565	Good Characters		Brown Eyes	Black Hair	Male Characters
1845	Neutral Characters		White Eyes	Red Hair	Female Characters
1937	Neutral Characters		Brown Eyes	Black Hair	Male Characters
2033	Neutral Characters		Blue Eyes	Black Hair	Female Characters
2147	Good Characters		NaN	NaN	Male Characters
2148	Good Characters		Blue Eyes	Black Hair	Male Characters
2398	Good Characters		NaN	NaN	Male Characters
2399	Bad Characters		White Eyes	Blue Hair	Female Characters
2549	Bad Characters		Red Eyes	NaN	Male Characters
2727	Good Characters		Red Eyes	No Hair	Male Characters
2728	Good Characters		Brown Eyes	Black Hair	Female Characters
2909	Good Characters		Brown Eyes	Brown Hair	Female Characters
2910	Good Characters		Purple Eyes	Blond Hair	Female Characters
2911	Good Characters		Brown Eyes	Brown Hair	Female Characters
...		...	...	...	...
16346		NaN	NaN	NaN	Male Characters
16347		NaN	NaN	NaN	NaN
16348		NaN	NaN	No Hair	NaN
16349	Good Characters		White Eyes	NaN	Male Characters
16350		NaN	NaN	NaN	NaN
16351	Bad Characters		NaN	Black Hair	Male Characters
16352		NaN	NaN	NaN	Male Characters
16353		NaN	NaN	NaN	Male Characters
16354	Good Characters		NaN	NaN	Male Characters
16355	Neutral Characters		NaN	NaN	Male Characters
16356	Bad Characters		NaN	No Hair	Male Characters
16357	Neutral Characters		NaN	NaN	Male Characters
16358	Bad Characters		NaN	Blond Hair	Female Characters
16359	Bad Characters		Black Eyes	No Hair	Male Characters
16360	Good Characters		Black Eyes	Grey Hair	Female Characters
16361	Neutral Characters		NaN	NaN	Female Characters
16362	Bad Characters		Red Eyes	NaN	Female Characters
16363	Good Characters		Black Eyes	Bald	Male Characters
16364	Bad Characters		Hazel Eyes	Bald	Male Characters
16365	Good Characters		NaN	NaN	Male Characters
16366	Neutral Characters		Brown Eyes	Black Hair	Female Characters
16367	Good Characters		Hazel Eyes	Bald	Male Characters
16368	Neutral Characters		NaN	Brown Hair	Female Characters
16369	Good Characters		Blue Eyes	Black Hair	Female Characters
16370	Neutral Characters		NaN	NaN	NaN
16371	Bad Characters		Green Eyes	No Hair	Male Characters
16372	Good Characters		Blue Eyes	Bald	Male Characters
16373	Bad Characters		Black Eyes	Bald	Male Characters

16374	Neutral Characters	NaN	NaN	Male Characters
16375	Bad Characters	NaN	NaN	NaN

	GSM	ALIVE	APPEARANCES \
12	NaN	Living Characters	1528.0
38	NaN	Living Characters	850.0
80	NaN	Deceased Characters	429.0
114	NaN	Living Characters	327.0
259	NaN	Living Characters	158.0
310	NaN	Living Characters	134.0
413	NaN	Living Characters	100.0
683	NaN	Living Characters	55.0
789	NaN	Deceased Characters	48.0
854	NaN	Living Characters	44.0
997	NaN	Living Characters	38.0
1118	NaN	Living Characters	34.0
1158	NaN	Deceased Characters	33.0
1316	NaN	Living Characters	29.0
1454	NaN	Deceased Characters	26.0
1564	NaN	Deceased Characters	24.0
1565	NaN	Living Characters	24.0
1845	NaN	Living Characters	20.0
1937	NaN	Living Characters	19.0
2033	NaN	Deceased Characters	18.0
2147	NaN	Living Characters	17.0
2148	NaN	Living Characters	17.0
2398	NaN	Living Characters	15.0
2399	NaN	Living Characters	15.0
2549	NaN	Living Characters	14.0
2727	NaN	Living Characters	13.0
2728	NaN	Living Characters	13.0
2909	NaN	Living Characters	12.0
2910	NaN	Living Characters	12.0
2911	NaN	Living Characters	12.0
...	...	...	...
16346	NaN	Deceased Characters	NaN
16347	NaN	NaN	NaN
16348	NaN	Living Characters	NaN
16349	NaN	Living Characters	NaN
16350	NaN	Living Characters	NaN
16351	NaN	Living Characters	NaN
16352	NaN	Living Characters	NaN
16353	NaN	Living Characters	NaN
16354	NaN	Deceased Characters	NaN
16355	NaN	Living Characters	NaN
16356	NaN	Living Characters	NaN
16357	NaN	Living Characters	NaN
16358	NaN	Living Characters	NaN

16359		NaN	Living Characters	NaN
16360		NaN	Living Characters	NaN
16361		NaN	Deceased Characters	NaN
16362		NaN	Living Characters	NaN
16363	Homosexual Characters		Living Characters	NaN
16364		NaN	Living Characters	NaN
16365		NaN	Deceased Characters	NaN
16366		NaN	Living Characters	NaN
16367	Homosexual Characters		Living Characters	NaN
16368		NaN	Living Characters	NaN
16369		NaN	Living Characters	NaN
16370		NaN	Living Characters	NaN
16371		NaN	Living Characters	NaN
16372		NaN	Living Characters	NaN
16373		NaN	Living Characters	NaN
16374		NaN	Living Characters	NaN
16375		NaN	Living Characters	NaN

	FIRST APPEARANCE	Year
12	NaN	NaN
38	NaN	NaN
80	NaN	NaN
114	NaN	NaN
259	NaN	NaN
310	NaN	NaN
413	NaN	NaN
683	NaN	NaN
789	NaN	NaN
854	NaN	NaN
997	NaN	NaN
1118	NaN	NaN
1158	NaN	NaN
1316	NaN	NaN
1454	NaN	NaN
1564	NaN	NaN
1565	NaN	NaN
1845	NaN	NaN
1937	NaN	NaN
2033	NaN	NaN
2147	NaN	NaN
2148	NaN	NaN
2398	NaN	NaN
2399	NaN	NaN
2549	NaN	NaN
2727	NaN	NaN
2728	NaN	NaN
2909	NaN	NaN
2910	NaN	NaN

2911	NaN	NaN
...	...	...
16346	NaN	NaN
16347	NaN	NaN
16348	NaN	NaN
16349	NaN	NaN
16350	NaN	NaN
16351	NaN	NaN
16352	NaN	NaN
16353	NaN	NaN
16354	NaN	NaN
16355	NaN	NaN
16356	NaN	NaN
16357	NaN	NaN
16358	NaN	NaN
16359	NaN	NaN
16360	NaN	NaN
16361	NaN	NaN
16362	NaN	NaN
16363	NaN	NaN
16364	NaN	NaN
16365	NaN	NaN
16366	NaN	NaN
16367	NaN	NaN
16368	NaN	NaN
16369	NaN	NaN
16370	NaN	NaN
16371	NaN	NaN
16372	NaN	NaN
16373	NaN	NaN
16374	NaN	NaN
16375	NaN	NaN

[815 rows x 13 columns]

```
In [82]: data[data['Year'].isnull()]
```

```
#
```

```
flt_index = data[data['Year'].isnull()].index
flt_index
```

```
Out[82]: Int64Index([ 12,   38,   80,  114,  259,  310,  413,  683,  789,
                    854,
                    ...,
                    16366, 16367, 16368, 16369, 16370, 16371, 16372, 16373, 16374,
                    16375],
                    dtype='int64', length=815)
```

```
In [83]: #
```

```
data_num[data_num.index.isin(flt_index)]['Year']
```

```

Out [83] : 12      NaN
           38      NaN
           80      NaN
           114     NaN
           259     NaN
           310     NaN
           413     NaN
           683     NaN
           789     NaN
           854     NaN
           997     NaN
           1118    NaN
           1158    NaN
           1316    NaN
           1454    NaN
           1564    NaN
           1565    NaN
           1845    NaN
           1937    NaN
           2033    NaN
           2147    NaN
           2148    NaN
           2398    NaN
           2399    NaN
           2549    NaN
           2727    NaN
           2728    NaN
           2909    NaN
           2910    NaN
           2911    NaN
           ..
           16346   NaN
           16347   NaN
           16348   NaN
           16349   NaN
           16350   NaN
           16351   NaN
           16352   NaN
           16353   NaN
           16354   NaN
           16355   NaN
           16356   NaN
           16357   NaN
           16358   NaN
           16359   NaN
           16360   NaN
           16361   NaN
           16362   NaN

```



```

16363    NaN
16364    NaN
16365    NaN
16366    NaN
16367    NaN
16368    NaN
16369    NaN
16370    NaN
16371    NaN
16372    NaN
16373    NaN
16374    NaN
16375    NaN
Name: Year, Length: 815, dtype: float64

```

```

In [86]: from sklearn.impute import SimpleImputer
         from sklearn.impute import MissingIndicator

```

```

In [96]: data[['Year']].describe()

```

```

Out[96]:
          Year
count  15561.000000
mean    1984.951803
std      19.663571
min     1939.000000
25%     1974.000000
50%     1990.000000
75%     2000.000000
max     2013.000000

```

```

In [105]: #
          #
          num_cols = []
          for col in data.columns:
              #
              temp_null_count = data[data[col].isnull()].shape[0]
              dt = str(data[col].dtype)
              if temp_null_count>0 and (dt=='float64' or dt=='int64'):
                  num_cols.append(col)
                  print(' {}.  {}.  {}'.format(col, dt, temp_null_count))

```

```

APPEARANCES.  float64.  1096.
Year.  float64.  815.

```

```

In [107]: #          Year
          data[data['Year'].isnull()]
          #
          flt_index = data[data['Year'].isnull()].index
          flt_index

```

```
Out[107]: Int64Index([ 12,    38,    80,   114,   259,   310,   413,   683,   789,
                     854,
                     ...,
                     16366, 16367, 16368, 16369, 16370, 16371, 16372, 16373, 16374,
                     16375],
                     dtype='int64', length=815)
```

```
In [116]: for rows in flt_index:
           data.Year[rows]=data.Year.median()
```

c:\program files\python36\lib\site-packages\ipykernel\_launcher.py:2: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html>

### 1.2.2.

```
In [208]: #
           #
           cat_cols = []
           for col in data.columns:
               #
               temp_null_count = data[data[col].isnull()].shape[0]
               dt = str(data[col].dtype)
               if temp_null_count>0 and (dt=='object'):
                   cat_cols.append(col)
                   temp_perc = round((temp_null_count / data.shape[0]) * 100.0, 2)
                   print(' {}.   {}.   {}, {}%.'.format(col, dt, temp_null_count,temp_perc))
```

```
ID.   object.   3770, 23.02%.
ALIGN. object.   2812, 17.17%.
EYE.   object.   9767, 59.64%.
HAIR.   object.   4264, 26.04%.
SEX.   object.   854, 5.21%.
GSM.   object.  16286, 99.45%.
ALIVE.  object.    3, 0.02%.
FIRST APPEARANCE.  object.   815, 4.98%.
```

```
In [209]: data_num = data[cat_cols]
           data_num
```

```
Out[209]:
```

	ID	ALIGN	EYE	HAIR \
0	Secret Identity	Good Characters	Hazel Eyes	Brown Hair
1	Public Identity	Good Characters	Blue Eyes	White Hair
2	Public Identity	Neutral Characters	Blue Eyes	Black Hair
3	Public Identity	Good Characters	Blue Eyes	Black Hair

4	No Dual Identity	Good Characters	Blue Eyes	Blond Hair
5	Public Identity	Good Characters	Blue Eyes	No Hair
6	Public Identity	Good Characters	Brown Eyes	Brown Hair
7	Public Identity	Good Characters	Brown Eyes	Brown Hair
8	Public Identity	Neutral Characters	Brown Eyes	Brown Hair
9	Public Identity	Good Characters	Blue Eyes	Blond Hair
10	Public Identity	Good Characters	Blue Eyes	Blue Hair
11	Public Identity	Good Characters	Blue Eyes	Blond Hair
12	No Dual Identity	Neutral Characters	Green Eyes	Black Hair
13	Public Identity	Good Characters	Blue Eyes	White Hair
14	Public Identity	Good Characters	Blue Eyes	Blond Hair
15	Public Identity	Good Characters	Blue Eyes	Red Hair
16	Public Identity	Good Characters	Grey Eyes	Black Hair
17	No Dual Identity	Good Characters	Green Eyes	Red Hair
18	No Dual Identity	Neutral Characters	Blue Eyes	Black Hair
19	Secret Identity	Good Characters	Brown Eyes	Brown Hair
20	Public Identity	Good Characters	Blue Eyes	Blond Hair
21	Public Identity	Good Characters	Blue Eyes	Bald
22	Public Identity	Good Characters	Blue Eyes	Blond Hair
23	Secret Identity	Good Characters	Blue Eyes	Black Hair
24	Public Identity	Good Characters	Green Eyes	Brown Hair
25	No Dual Identity	Neutral Characters	Brown Eyes	Brown Hair
26	Public Identity	Good Characters	Blue Eyes	Auburn Hair
27	Public Identity	Good Characters	Green Eyes	Red Hair
28	Public Identity	Good Characters	Green Eyes	Red Hair
29	Secret Identity	Good Characters	Yellow Eyes	Blue Hair
...	...	...	...	...
16346	NaN	NaN	NaN	NaN
16347	NaN	NaN	NaN	NaN
16348	Secret Identity	NaN	NaN	No Hair
16349	Public Identity	Good Characters	White Eyes	NaN
16350	NaN	NaN	NaN	NaN
16351	No Dual Identity	Bad Characters	NaN	Black Hair
16352	NaN	NaN	NaN	NaN
16353	NaN	NaN	NaN	NaN
16354	NaN	Good Characters	NaN	NaN
16355	Public Identity	Neutral Characters	NaN	NaN
16356	Public Identity	Bad Characters	NaN	No Hair
16357	Secret Identity	Neutral Characters	NaN	NaN
16358	Secret Identity	Bad Characters	NaN	Blond Hair
16359	No Dual Identity	Bad Characters	Black Eyes	No Hair
16360	No Dual Identity	Good Characters	Black Eyes	Grey Hair
16361	No Dual Identity	Neutral Characters	NaN	NaN
16362	Public Identity	Bad Characters	Red Eyes	NaN
16363	Public Identity	Good Characters	Black Eyes	Bald
16364	Secret Identity	Bad Characters	Hazel Eyes	Bald
16365	NaN	Good Characters	NaN	NaN
16366	Public Identity	Neutral Characters	Brown Eyes	Black Hair

16367	Public Identity	Good Characters	Hazel Eyes	Bald
16368	Public Identity	Neutral Characters	NaN	Brown Hair
16369	Public Identity	Good Characters	Blue Eyes	Black Hair
16370	NaN	Neutral Characters	NaN	NaN
16371	No Dual Identity	Bad Characters	Green Eyes	No Hair
16372	No Dual Identity	Good Characters	Blue Eyes	Bald
16373	Secret Identity	Bad Characters	Black Eyes	Bald
16374	Secret Identity	Neutral Characters	NaN	NaN
16375	NaN	Bad Characters	NaN	NaN

	SEX		GSM	ALIVE \
0	Male Characters		NaN	Living Characters
1	Male Characters		NaN	Living Characters
2	Male Characters		NaN	Living Characters
3	Male Characters		NaN	Living Characters
4	Male Characters		NaN	Living Characters
5	Male Characters		NaN	Living Characters
6	Male Characters		NaN	Living Characters
7	Male Characters		NaN	Living Characters
8	Male Characters		NaN	Living Characters
9	Male Characters		NaN	Living Characters
10	Male Characters		NaN	Living Characters
11	Female Characters		NaN	Living Characters
12	Male Characters		NaN	Living Characters
13	Female Characters		NaN	Living Characters
14	Male Characters		NaN	Living Characters
15	Male Characters		NaN	Living Characters
16	Male Characters		NaN	Living Characters
17	Female Characters		NaN	Living Characters
18	Male Characters		NaN	Living Characters
19	Male Characters		NaN	Living Characters
20	Male Characters		NaN	Living Characters
21	Male Characters		NaN	Deceased Characters
22	Male Characters		NaN	Living Characters
23	Male Characters		NaN	Living Characters
24	Female Characters		NaN	Living Characters
25	Male Characters		NaN	Living Characters
26	Female Characters		NaN	Living Characters
27	Female Characters		NaN	Deceased Characters
28	Female Characters	Bisexual Characters		Living Characters
29	Male Characters		NaN	Living Characters
...	...		...	...
16346	Male Characters		NaN	Deceased Characters
16347	NaN		NaN	NaN
16348	NaN		NaN	Living Characters
16349	Male Characters		NaN	Living Characters
16350	NaN		NaN	Living Characters
16351	Male Characters		NaN	Living Characters

16352	Male Characters	NaN	Living Characters
16353	Male Characters	NaN	Living Characters
16354	Male Characters	NaN	Deceased Characters
16355	Male Characters	NaN	Living Characters
16356	Male Characters	NaN	Living Characters
16357	Male Characters	NaN	Living Characters
16358	Female Characters	NaN	Living Characters
16359	Male Characters	NaN	Living Characters
16360	Female Characters	NaN	Living Characters
16361	Female Characters	NaN	Deceased Characters
16362	Female Characters	NaN	Living Characters
16363	Male Characters	Homosexual Characters	Living Characters
16364	Male Characters	NaN	Living Characters
16365	Male Characters	NaN	Deceased Characters
16366	Female Characters	NaN	Living Characters
16367	Male Characters	Homosexual Characters	Living Characters
16368	Female Characters	NaN	Living Characters
16369	Female Characters	NaN	Living Characters
16370	NaN	NaN	Living Characters
16371	Male Characters	NaN	Living Characters
16372	Male Characters	NaN	Living Characters
16373	Male Characters	NaN	Living Characters
16374	Male Characters	NaN	Living Characters
16375	NaN	NaN	Living Characters

#### FIRST APPEARANCE

0	Aug-62
1	Mar-41
2	Oct-74
3	Mar-63
4	Nov-50
5	Nov-61
6	Nov-61
7	May-62
8	Sep-63
9	Nov-61
10	Sep-63
11	Nov-61
12	NaN
13	May-75
14	Sep-64
15	Apr-64
16	Jul-63
17	Jun-65
18	Mar-63
19	Sep-63
20	Jan-62
21	Sep-63

22	Sep-63
23	May-75
24	Mar-64
25	May-63
26	Jun-63
27	Sep-63
28	Apr-64
29	May-75
...	...
16346	NaN
16347	NaN
16348	NaN
16349	NaN
16350	NaN
16351	NaN
16352	NaN
16353	NaN
16354	NaN
16355	NaN
16356	NaN
16357	NaN
16358	NaN
16359	NaN
16360	NaN
16361	NaN
16362	NaN
16363	NaN
16364	NaN
16365	NaN
16366	NaN
16367	NaN
16368	NaN
16369	NaN
16370	NaN
16371	NaN
16372	NaN
16373	NaN
16374	NaN
16375	NaN

[16376 rows x 8 columns]

```
In [198]: #      Year
data[data['HAIR'].isnull()]
#
flt_index = data[data['HAIR'].isnull()].index
flt_index
```

```
Out[198]: Int64Index([], dtype='int64')
```

```
In [212]: data[['HAIR']].describe()
```

```
Out[212]:
```

	HAIR
count	16376
unique	25
top	Purple Hair
freq	4311

```
In [211]: MaxPassEmbarked = data.groupby('HAIR').count()['name']
data.HAIR[data.HAIR.isnull()] = MaxPassEmbarked[MaxPassEmbarked == MaxPassEmbarked.max()]

data[data[col].isnull()].shape[0]
```

c:\program files\python36\lib\site-packages\ipykernel\_launcher.py:2: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html>

```
Out[211]: 815
```

```
In [182]: data.ALIGN.replace({'Good Characters':'1','Bad Characters':'0','Neutral Characters':'2'})
data.head(50)
```

```
Out[182]:
```

	page_id	name \
0	1678	Spider-Man (Peter Parker)
1	7139	Captain America (Steven Rogers)
2	64786	Wolverine (James \"Logan\" Howlett)
3	1868	Iron Man (Anthony \"Tony\" Stark)
4	2460	Thor (Thor Odinson)
5	2458	Benjamin Grimm (Earth-616)
6	2166	Reed Richards (Earth-616)
7	1833	Hulk (Robert Bruce Banner)
8	29481	Scott Summers (Earth-616)
9	1837	Jonathan Storm (Earth-616)
10	15725	Henry McCoy (Earth-616)
11	1863	Susan Storm (Earth-616)
12	7823	Namor McKenzie (Earth-616)
13	2614	Ororo Munroe (Earth-616)
14	1803	Clinton Barton (Earth-616)
15	1396	Matthew Murdock (Earth-616)
16	55534	Stephen Strange (Earth-616)
17	1978	Mary Jane Watson (Earth-616)
18	1872	John Jonah Jameson (Earth-616)
19	35350	Robert Drake (Earth-616)
20	1557	Henry Pym (Earth-616)

21	65255	Charles Xavier	(Earth-616)
22	1073	Warren Worthington III	(Earth-616)
23	1346	Piotr Rasputin	(Earth-616)
24	2512	Wanda Maximoff	(Earth-616)
25	1671	Nicholas Fury	(Earth-616)
26	1976	Janet van Dyne	(Earth-616)
27	65219	Jean Grey	(Earth-616)
28	6545	Natalia Romanova	(Earth-616)
29	2223	Kurt Wagner	(Earth-616)
30	2414	Vision	(Earth-616)
31	8650	May Reilly	(Earth-616)
32	2527	Katherine Pryde	(Earth-616)
33	1970	Carol Danvers	(Earth-616)
34	37690	Jennifer Walters	(Earth-616)
35	3765	Emma Frost	(Earth-616)
36	2312	Frank Castle	(Earth-616)
37	1265	Luke Cage	(Earth-616)
38	1677	Rogue (Anna Marie)	(Earth-616)
39	25222	Conan	(Earth-616)
40	6537	Joseph Robertson	(Earth-616)
41	1592	Pietro Maximoff	(Earth-616)
42	1818	Hercules	(Earth-616)
43	1448	Victor von Doom	(Earth-616)
44	2081	Max Eisenhardt	(Earth-616)
45	2307	Elizabeth Braddock	(Earth-616)
46	2556	Norrin Radd	(Earth-616)
47	1740	Norman Osborn	(Earth-616)
48	8651	Eugene Thompson	(Earth-616)
49	2009	Simon Williams	(Earth-616)

	urlslug	ID	ALIGN	\
0	\\/Spider-Man_(Peter_Parker)	Secret	Identity	1
1	\\/Captain_America_(Steven_Rogers)	Public	Identity	1
2	\\/Wolverine_(James_%22Logan%22_Howlett)	Public	Identity	2
3	\\/Iron_Man_(Anthony_%22Tony%22_Stark)	Public	Identity	1
4	\\/Thor_(Thor_Odinson)	No Dual	Identity	1
5	\\/Benjamin_Grimm_(Earth-616)	Public	Identity	1
6	\\/Reed_Richards_(Earth-616)	Public	Identity	1
7	\\/Hulk_(Robert_Bruce_Banner)	Public	Identity	1
8	\\/Scott_Summers_(Earth-616)	Public	Identity	2
9	\\/Jonathan_Storm_(Earth-616)	Public	Identity	1
10	\\/Henry_McCoy_(Earth-616)	Public	Identity	1
11	\\/Susan_Storm_(Earth-616)	Public	Identity	1
12	\\/Namor_McKenzie_(Earth-616)	No Dual	Identity	2
13	\\/Ororo_Munroe_(Earth-616)	Public	Identity	1
14	\\/Clinton_Barton_(Earth-616)	Public	Identity	1
15	\\/Matthew_Murdock_(Earth-616)	Public	Identity	1
16	\\/Stephen_Strange_(Earth-616)	Public	Identity	1



17	\\Mary_Jane_Watson_(Earth-616)	No Dual Identity	1
18	\\John_Jonah_Jameson_(Earth-616)	No Dual Identity	2
19	\\Robert_Drake_(Earth-616)	Secret Identity	1
20	\\Henry_Pym_(Earth-616)	Public Identity	1
21	\\Charles_Xavier_(Earth-616)	Public Identity	1
22	\\Warren_Worthington_III_(Earth-616)	Public Identity	1
23	\\Piotr_Rasputin_(Earth-616)	Secret Identity	1
24	\\Wanda_Maximoff_(Earth-616)	Public Identity	1
25	\\Nicholas_Fury_(Earth-616)	No Dual Identity	2
26	\\Janet_van_Dyne_(Earth-616)	Public Identity	1
27	\\Jean_Grey_(Earth-616)	Public Identity	1
28	\\Natalia_Romanova_(Earth-616)	Public Identity	1
29	\\Kurt_Wagner_(Earth-616)	Secret Identity	1
30	\\Vision_(Earth-616)	Secret Identity	1
31	\\May_Reilly_(Earth-616)	No Dual Identity	1
32	\\Katherine_Pryde_(Earth-616)	Secret Identity	1
33	\\Carol_Danvers_(Earth-616)	Public Identity	1
34	\\Jennifer_Walters_(Earth-616)	Public Identity	1
35	\\Emma_Frost_(Earth-616)	Public Identity	2
36	\\Frank_Castle_(Earth-616)	Public Identity	2
37	\\Luke_Cage_(Earth-616)	No Dual Identity	1
38	\\Rogue_(Anna_Marie)_(Earth-616)	Secret Identity	1
39	\\Conan_(Earth-616)	No Dual Identity	2
40	\\Joseph_Robertson_(Earth-616)	No Dual Identity	1
41	\\Pietro_Maximoff_(Earth-616)	Public Identity	1
42	\\Hercules_(Earth-616)	No Dual Identity	1
43	\\Victor_von_Doom_(Earth-616)	Public Identity	0
44	\\Max_Eisenhardt_(Earth-616)	Public Identity	2
45	\\Elizabeth_Braddock_(Earth-616)	Secret Identity	2
46	\\Norrin_Radd_(Earth-616)	Public Identity	1
47	\\Norman_Osborn_(Earth-616)	Public Identity	0
48	\\Eugene_Thompson_(Earth-616)	Secret Identity	1
49	\\Simon_Williams_(Earth-616)	Public Identity	1

	EYE	HAIR	SEX \
0	Hazel Eyes	Brown Hair	Male Characters
1	Blue Eyes	White Hair	Male Characters
2	Blue Eyes	Black Hair	Male Characters
3	Blue Eyes	Black Hair	Male Characters
4	Blue Eyes	Blond Hair	Male Characters
5	Blue Eyes	No Hair	Male Characters
6	Brown Eyes	Brown Hair	Male Characters
7	Brown Eyes	Brown Hair	Male Characters
8	Brown Eyes	Brown Hair	Male Characters
9	Blue Eyes	Blond Hair	Male Characters
10	Blue Eyes	Blue Hair	Male Characters
11	Blue Eyes	Blond Hair	Female Characters
12	Green Eyes	Black Hair	Male Characters

13	Blue Eyes	White Hair	Female Characters
14	Blue Eyes	Blond Hair	Male Characters
15	Blue Eyes	Red Hair	Male Characters
16	Grey Eyes	Black Hair	Male Characters
17	Green Eyes	Red Hair	Female Characters
18	Blue Eyes	Black Hair	Male Characters
19	Brown Eyes	Brown Hair	Male Characters
20	Blue Eyes	Blond Hair	Male Characters
21	Blue Eyes	Bald	Male Characters
22	Blue Eyes	Blond Hair	Male Characters
23	Blue Eyes	Black Hair	Male Characters
24	Green Eyes	Brown Hair	Female Characters
25	Brown Eyes	Brown Hair	Male Characters
26	Blue Eyes	Auburn Hair	Female Characters
27	Green Eyes	Red Hair	Female Characters
28	Green Eyes	Red Hair	Female Characters
29	Yellow Eyes	Blue Hair	Male Characters
30	Gold Eyes	No Hair	Male Characters
31	Blue Eyes	Grey Hair	Female Characters
32	Hazel Eyes	Brown Hair	Female Characters
33	Blue Eyes	Blond Hair	Female Characters
34	Green Eyes	Brown Hair	Female Characters
35	Blue Eyes	Brown Hair	Female Characters
36	Blue Eyes	Black Hair	Male Characters
37	Brown Eyes	Black Hair	Male Characters
38	Green Eyes	Auburn Hair	Female Characters
39	Blue Eyes	Black Hair	Male Characters
40	Brown Eyes	White Hair	Male Characters
41	Blue Eyes	Silver Hair	Male Characters
42	Blue Eyes	Brown Hair	Male Characters
43	Brown Eyes	Brown Hair	Male Characters
44	Grey Eyes	Bald	Male Characters
45	Blue Eyes	Purple Hair	Female Characters
46	Blue Eyes	Bald	Male Characters
47	Blue Eyes	Auburn Hair	Male Characters
48	Blue Eyes	Strawberry Blond Hair	Male Characters
49	Red Eyes	Grey Hair	Male Characters

	GSM	ALIVE	APPEARANCES	FIRST APPEARANCE	\
0	NaN	Living Characters	4043.0	Aug-62	
1	NaN	Living Characters	3360.0	Mar-41	
2	NaN	Living Characters	3061.0	Oct-74	
3	NaN	Living Characters	2961.0	Mar-63	
4	NaN	Living Characters	2258.0	Nov-50	
5	NaN	Living Characters	2255.0	Nov-61	
6	NaN	Living Characters	2072.0	Nov-61	
7	NaN	Living Characters	2017.0	May-62	
8	NaN	Living Characters	1955.0	Sep-63	

9		NaN	Living Characters	1934.0	Nov-61
10		NaN	Living Characters	1825.0	Sep-63
11		NaN	Living Characters	1713.0	Nov-61
12		NaN	Living Characters	1528.0	NaN
13		NaN	Living Characters	1512.0	May-75
14		NaN	Living Characters	1394.0	Sep-64
15		NaN	Living Characters	1338.0	Apr-64
16		NaN	Living Characters	1307.0	Jul-63
17		NaN	Living Characters	1304.0	Jun-65
18		NaN	Living Characters	1266.0	Mar-63
19		NaN	Living Characters	1265.0	Sep-63
20		NaN	Living Characters	1237.0	Jan-62
21		NaN	Deceased Characters	1233.0	Sep-63
22		NaN	Living Characters	1230.0	Sep-63
23		NaN	Living Characters	1162.0	May-75
24		NaN	Living Characters	1161.0	Mar-64
25		NaN	Living Characters	1137.0	May-63
26		NaN	Living Characters	1120.0	Jun-63
27		NaN	Deceased Characters	1107.0	Sep-63
28	Bisexual Characters		Living Characters	1050.0	Apr-64
29		NaN	Living Characters	1047.0	May-75
30		NaN	Living Characters	1007.0	Oct-68
31		NaN	Living Characters	1000.0	Aug-62
32		NaN	Living Characters	886.0	Jan-80
33		NaN	Living Characters	881.0	Mar-68
34		NaN	Living Characters	881.0	Feb-80
35		NaN	Living Characters	880.0	Jan-80
36		NaN	Living Characters	878.0	Feb-74
37		NaN	Living Characters	856.0	Jun-72
38		NaN	Living Characters	850.0	NaN
39		NaN	Living Characters	787.0	Oct-70
40		NaN	Living Characters	752.0	Aug-67
41		NaN	Living Characters	740.0	Mar-64
42	Bisexual Characters		Living Characters	725.0	Oct-65
43		NaN	Living Characters	721.0	Jul-62
44		NaN	Living Characters	709.0	Sep-63
45	Bisexual Characters		Living Characters	708.0	Dec-76
46		NaN	Living Characters	696.0	Mar-66
47		NaN	Living Characters	692.0	Jul-64
48		NaN	Living Characters	689.0	Aug-62
49		NaN	Living Characters	686.0	Oct-64

	Year
0	1962.0
1	1941.0
2	1974.0
3	1963.0
4	1950.0

```
5  1961.0
6  1961.0
7  1962.0
8  1963.0
9  1961.0
10 1963.0
11 1961.0
12   NaN
13 1975.0
14 1964.0
15 1964.0
16 1963.0
17 1965.0
18 1963.0
19 1963.0
20 1962.0
21 1963.0
22 1963.0
23 1975.0
24 1964.0
25 1963.0
26 1963.0
27 1963.0
28 1964.0
29 1975.0
30 1968.0
31 1962.0
32 1980.0
33 1968.0
34 1980.0
35 1980.0
36 1974.0
37 1972.0
38   NaN
39 1970.0
40 1967.0
41 1964.0
42 1965.0
43 1962.0
44 1963.0
45 1976.0
46 1966.0
47 1964.0
48 1962.0
49 1964.0
```

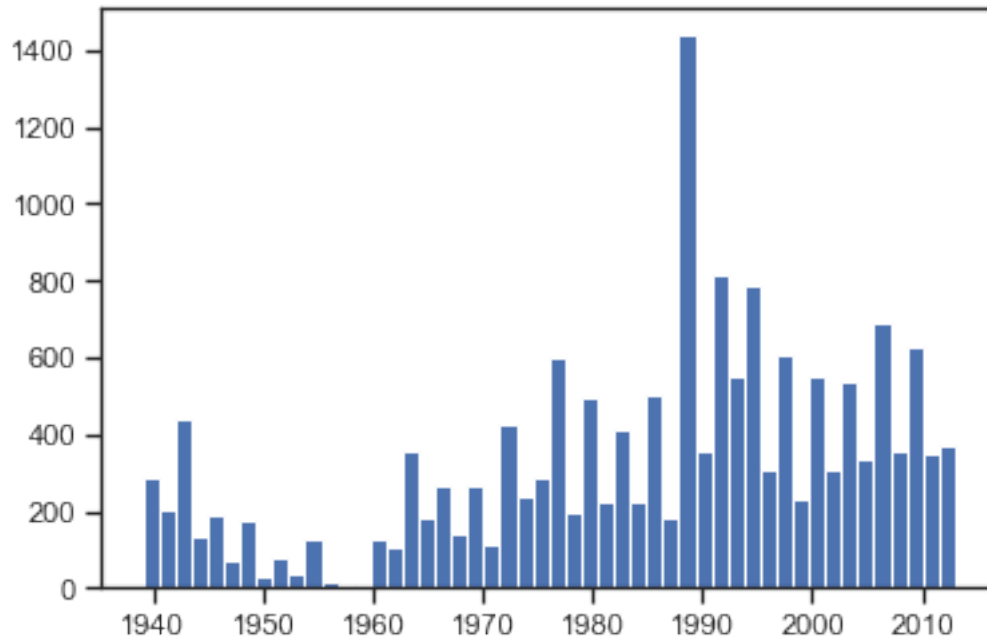
```
In [58]: from sklearn.preprocessing import MinMaxScaler, StandardScaler, Normalizer
```

```
In [124]: sc1 = MinMaxScaler()
```

```

sc1_data = sc1.fit_transform(data[['Year']])
plt.hist(data['Year'], 50)
plt.show()

```



Z-

```

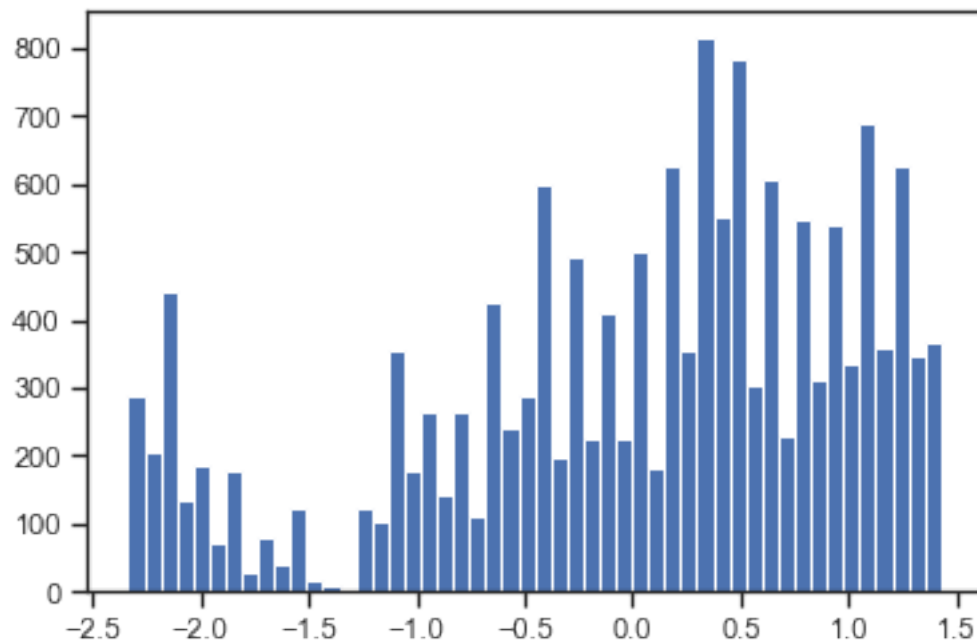
In [53]: sc2 = StandardScaler()
sc2_data = sc2.fit_transform(data[['Year']])
plt.hist(sc2_data, 50)
plt.show()

```

```

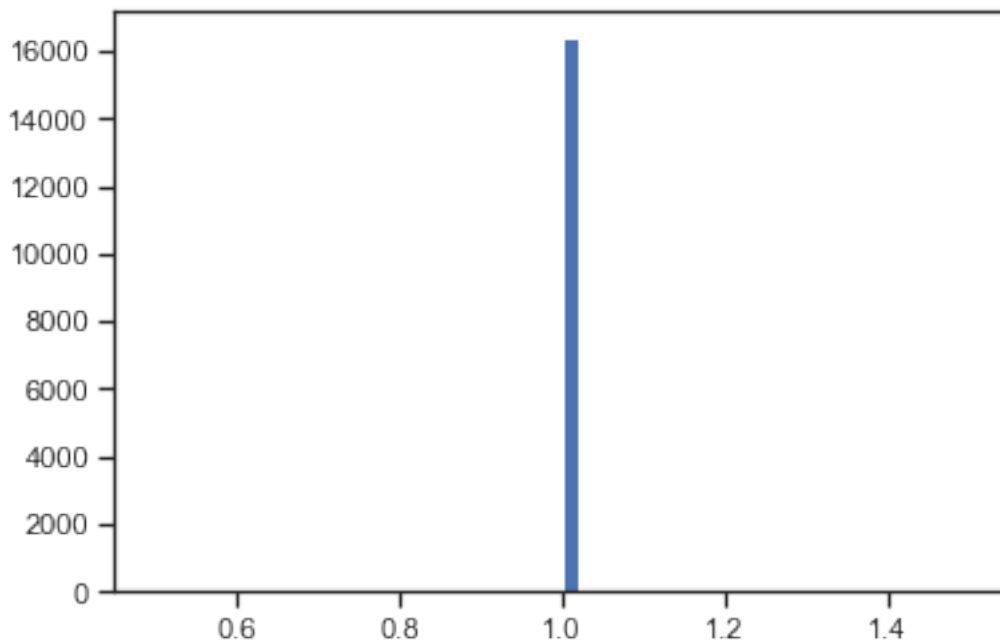
c:\program files\python36\lib\site-packages\numpy\lib\histograms.py:824: RuntimeWarning: invalid value encountered in less
  keep = (tmp_a >= first_edge)
c:\program files\python36\lib\site-packages\numpy\lib\histograms.py:825: RuntimeWarning: invalid value encountered in less
  keep &= (tmp_a <= last_edge)

```



```
In [120]: sc3 = Normalizer()  
          sc3_data = sc3.fit_transform(data[['Year']])
```

```
In [121]: plt.hist(sc3_data, 50)  
          plt.show()
```



```
In [213]: data.head(17000)
```

```
Out[213]:
```

	page_id	name \
0	1678	Spider-Man (Peter Parker)
1	7139	Captain America (Steven Rogers)
2	64786	Wolverine (James \"Logan\" Howlett)
3	1868	Iron Man (Anthony \"Tony\" Stark)
4	2460	Thor (Thor Odinson)
5	2458	Benjamin Grimm (Earth-616)
6	2166	Reed Richards (Earth-616)
7	1833	Hulk (Robert Bruce Banner)
8	29481	Scott Summers (Earth-616)
9	1837	Jonathan Storm (Earth-616)
10	15725	Henry McCoy (Earth-616)
11	1863	Susan Storm (Earth-616)
12	7823	Namor McKenzie (Earth-616)
13	2614	Ororo Munroe (Earth-616)
14	1803	Clinton Barton (Earth-616)
15	1396	Matthew Murdock (Earth-616)
16	55534	Stephen Strange (Earth-616)
17	1978	Mary Jane Watson (Earth-616)
18	1872	John Jonah Jameson (Earth-616)
19	35350	Robert Drake (Earth-616)
20	1557	Henry Pym (Earth-616)
21	65255	Charles Xavier (Earth-616)
22	1073	Warren Worthington III (Earth-616)
23	1346	Piotr Rasputin (Earth-616)
24	2512	Wanda Maximoff (Earth-616)
25	1671	Nicholas Fury (Earth-616)
26	1976	Janet van Dyne (Earth-616)
27	65219	Jean Grey (Earth-616)
28	6545	Natalia Romanova (Earth-616)
29	2223	Kurt Wagner (Earth-616)
...	...	...
16346	42106	Thiazzzi (Earth-616)
16347	462671	TOR/test
16348	725312	Toxin (Luminals) (Earth-616)
16349	507239	Urizen Ul'var (Earth-616)
16350	40061	Valka (Earth-616)
16351	505759	Vayor (Earth-616)
16352	16569	Viridian (Earth-616)
16353	16296	William Burke (Earth-616)
16354	120833	William Falsworth (Earth-616)
16355	693077	William Shakespeare (Earth-616)
16356	473710	Yamata no Oricchi (Earth-616)

16357	670512	Zero G. Priestly (Legion Personality)	(Earth-616)
16358	12132	Zora Loftus	(Earth-616)
16359	727035	Agar	(Earth-616)
16360	703546	Ana (Natasha Romanoff's neighbor)	(Earth-616)
16361	713903	Dante's mother	(Earth-616)
16362	508693	Farbauti	(Earth-616)
16363	694577	Finch (Kate Bishop's neighbor)	(Earth-616)
16364	655292	Jack O'Lantern (Impostor)	(Earth-616)
16365	684262	K'thol	(Earth-616)
16366	643435	Karen (Hijack's girlfriend)	(Earth-616)
16367	694583	Marcus (Kate Bishop's neighbor)	(Earth-616)
16368	703892	Marcy (Offer's employee)	(Earth-616)
16369	660799	Melanie Kapoor	(Earth-616)
16370	674414	Phoenix's Shadow	(Earth-616)
16371	657508	Ru'ach	(Earth-616)
16372	665474	Thane (Thanos' son)	(Earth-616)
16373	695217	Tinkerer (Skrull)	(Earth-616)
16374	708811	TK421 (Spiderling)	(Earth-616)
16375	673702	Yologarch	(Earth-616)

	urlslug	ID	\
0	\\Spider-Man_(Peter_Parker)	Secret Identity	
1	\\Captain_America_(Steven_Rogers)	Public Identity	
2	\\Wolverine_(James_%22Logan%22_Howlett)	Public Identity	
3	\\Iron_Man_(Anthony_%22Tony%22_Stark)	Public Identity	
4	\\Thor_(Thor_Odinson)	No Dual Identity	
5	\\Benjamin_Grimm_(Earth-616)	Public Identity	
6	\\Reed_Richards_(Earth-616)	Public Identity	
7	\\Hulk_(Robert_Bruce_Banner)	Public Identity	
8	\\Scott_Summers_(Earth-616)	Public Identity	
9	\\Jonathan_Storm_(Earth-616)	Public Identity	
10	\\Henry_McCoy_(Earth-616)	Public Identity	
11	\\Susan_Storm_(Earth-616)	Public Identity	
12	\\Namor_McKenzie_(Earth-616)	No Dual Identity	
13	\\Ororo_Munroe_(Earth-616)	Public Identity	
14	\\Clinton_Barton_(Earth-616)	Public Identity	
15	\\Matthew_Murdock_(Earth-616)	Public Identity	
16	\\Stephen_Strange_(Earth-616)	Public Identity	
17	\\Mary_Jane_Watson_(Earth-616)	No Dual Identity	
18	\\John_Jonah_Jameson_(Earth-616)	No Dual Identity	
19	\\Robert_Drake_(Earth-616)	Secret Identity	
20	\\Henry_Pym_(Earth-616)	Public Identity	
21	\\Charles_Xavier_(Earth-616)	Public Identity	
22	\\Warren_Worthington_III_(Earth-616)	Public Identity	
23	\\Piotr_Rasputin_(Earth-616)	Secret Identity	
24	\\Wanda_Maximoff_(Earth-616)	Public Identity	
25	\\Nicholas_Fury_(Earth-616)	No Dual Identity	
26	\\Janet_van_Dyne_(Earth-616)	Public Identity	



27	\\Jean_Grey_(Earth-616)	Public Identity
28	\\Natalia_Romanova_(Earth-616)	Public Identity
29	\\Kurt_Wagner_(Earth-616)	Secret Identity
...	...	...
16346	\\Thiazzi_(Earth-616)	NaN
16347	\\User:TOR\\test	NaN
16348	\\Toxin_(Luminals)_(Earth-616)	Secret Identity
16349	\\Urizen_U1%27var_(Earth-616)	Public Identity
16350	\\Valka_(Earth-616)	NaN
16351	\\Vayor_(Earth-616)	No Dual Identity
16352	\\Viridian_(Earth-616)	NaN
16353	\\William_Burke_(Earth-616)	NaN
16354	\\William_Falsworth_(Earth-616)	NaN
16355	\\William_Shakespeare_(Earth-616)	Public Identity
16356	\\Yamata_no_Orichi_(Earth-616)	Public Identity
16357	\\Zero_G._Priestly_(Legion_Personality)_(Earth...	Secret Identity
16358	\\Zora_Loftus_(Earth-616)	Secret Identity
16359	\\Agar_(Earth-616)	No Dual Identity
16360	\\Ana_(Natasha_Romanoff%27s_neighbor)_(Earth-616)	No Dual Identity
16361	\\Dante%27s_mother_(Earth-616)	No Dual Identity
16362	\\Farbauti_(Earth-616)	Public Identity
16363	\\Finch_(Kate_Bishop%27s_neighbor)_(Earth-616)	Public Identity
16364	\\Jack_0%27Lantern_(Impostor)_(Earth-616)	Secret Identity
16365	\\K%27thol_(Earth-616)	NaN
16366	\\Karen_(Hijack%27s_girlfriend)_(Earth-616)	Public Identity
16367	\\Marcus_(Kate_Bishop%27s_neighbor)_(Earth-616)	Public Identity
16368	\\Marcy_(Offer%27s_employee)_(Earth-616)	Public Identity
16369	\\Melanie_Kapoor_(Earth-616)	Public Identity
16370	\\Phoenix%27s_Shadow_(Earth-616)	NaN
16371	\\Ru%27ach_(Earth-616)	No Dual Identity
16372	\\Thane_(Thanos%27_son)_(Earth-616)	No Dual Identity
16373	\\Tinkerer_(Skrull)_(Earth-616)	Secret Identity
16374	\\TK421_(Spiderling)_(Earth-616)	Secret Identity
16375	\\Yologarch_(Earth-616)	NaN

	ALIGN	EYE	HAIR	SEX \
0	Good Characters	Hazel Eyes	Brown Hair	Male Characters
1	Good Characters	Blue Eyes	White Hair	Male Characters
2	Neutral Characters	Blue Eyes	Black Hair	Male Characters
3	Good Characters	Blue Eyes	Black Hair	Male Characters
4	Good Characters	Blue Eyes	Blond Hair	Male Characters
5	Good Characters	Blue Eyes	No Hair	Male Characters
6	Good Characters	Brown Eyes	Brown Hair	Male Characters
7	Good Characters	Brown Eyes	Brown Hair	Male Characters
8	Neutral Characters	Brown Eyes	Brown Hair	Male Characters
9	Good Characters	Blue Eyes	Blond Hair	Male Characters
10	Good Characters	Blue Eyes	Blue Hair	Male Characters
11	Good Characters	Blue Eyes	Blond Hair	Female Characters

12	Neutral Characters	Green Eyes	Black Hair	Male Characters
13	Good Characters	Blue Eyes	White Hair	Female Characters
14	Good Characters	Blue Eyes	Blond Hair	Male Characters
15	Good Characters	Blue Eyes	Red Hair	Male Characters
16	Good Characters	Grey Eyes	Black Hair	Male Characters
17	Good Characters	Green Eyes	Red Hair	Female Characters
18	Neutral Characters	Blue Eyes	Black Hair	Male Characters
19	Good Characters	Brown Eyes	Brown Hair	Male Characters
20	Good Characters	Blue Eyes	Blond Hair	Male Characters
21	Good Characters	Blue Eyes	Bald	Male Characters
22	Good Characters	Blue Eyes	Blond Hair	Male Characters
23	Good Characters	Blue Eyes	Black Hair	Male Characters
24	Good Characters	Green Eyes	Brown Hair	Female Characters
25	Neutral Characters	Brown Eyes	Brown Hair	Male Characters
26	Good Characters	Blue Eyes	Auburn Hair	Female Characters
27	Good Characters	Green Eyes	Red Hair	Female Characters
28	Good Characters	Green Eyes	Red Hair	Female Characters
29	Good Characters	Yellow Eyes	Blue Hair	Male Characters
...	...	...	...	...
16346	NaN	NaN	Purple Hair	Male Characters
16347	NaN	NaN	Purple Hair	NaN
16348	NaN	NaN	No Hair	NaN
16349	Good Characters	White Eyes	Purple Hair	Male Characters
16350	NaN	NaN	Purple Hair	NaN
16351	Bad Characters	NaN	Black Hair	Male Characters
16352	NaN	NaN	Purple Hair	Male Characters
16353	NaN	NaN	Purple Hair	Male Characters
16354	Good Characters	NaN	Purple Hair	Male Characters
16355	Neutral Characters	NaN	Purple Hair	Male Characters
16356	Bad Characters	NaN	No Hair	Male Characters
16357	Neutral Characters	NaN	Purple Hair	Male Characters
16358	Bad Characters	NaN	Blond Hair	Female Characters
16359	Bad Characters	Black Eyes	No Hair	Male Characters
16360	Good Characters	Black Eyes	Grey Hair	Female Characters
16361	Neutral Characters	NaN	Purple Hair	Female Characters
16362	Bad Characters	Red Eyes	Purple Hair	Female Characters
16363	Good Characters	Black Eyes	Bald	Male Characters
16364	Bad Characters	Hazel Eyes	Bald	Male Characters
16365	Good Characters	NaN	Purple Hair	Male Characters
16366	Neutral Characters	Brown Eyes	Black Hair	Female Characters
16367	Good Characters	Hazel Eyes	Bald	Male Characters
16368	Neutral Characters	NaN	Brown Hair	Female Characters
16369	Good Characters	Blue Eyes	Black Hair	Female Characters
16370	Neutral Characters	NaN	Purple Hair	NaN
16371	Bad Characters	Green Eyes	No Hair	Male Characters
16372	Good Characters	Blue Eyes	Bald	Male Characters
16373	Bad Characters	Black Eyes	Bald	Male Characters
16374	Neutral Characters	NaN	Purple Hair	Male Characters

16375	Bad Characters	NaN	Purple Hair	NaN
		GSM	ALIVE	APPEARANCES \
0		NaN	Living Characters	4043.0
1		NaN	Living Characters	3360.0
2		NaN	Living Characters	3061.0
3		NaN	Living Characters	2961.0
4		NaN	Living Characters	2258.0
5		NaN	Living Characters	2255.0
6		NaN	Living Characters	2072.0
7		NaN	Living Characters	2017.0
8		NaN	Living Characters	1955.0
9		NaN	Living Characters	1934.0
10		NaN	Living Characters	1825.0
11		NaN	Living Characters	1713.0
12		NaN	Living Characters	1528.0
13		NaN	Living Characters	1512.0
14		NaN	Living Characters	1394.0
15		NaN	Living Characters	1338.0
16		NaN	Living Characters	1307.0
17		NaN	Living Characters	1304.0
18		NaN	Living Characters	1266.0
19		NaN	Living Characters	1265.0
20		NaN	Living Characters	1237.0
21		NaN	Deceased Characters	1233.0
22		NaN	Living Characters	1230.0
23		NaN	Living Characters	1162.0
24		NaN	Living Characters	1161.0
25		NaN	Living Characters	1137.0
26		NaN	Living Characters	1120.0
27		NaN	Deceased Characters	1107.0
28	Bisexual Characters		Living Characters	1050.0
29		NaN	Living Characters	1047.0
...		...	...	...
16346		NaN	Deceased Characters	NaN
16347		NaN	NaN	NaN
16348		NaN	Living Characters	NaN
16349		NaN	Living Characters	NaN
16350		NaN	Living Characters	NaN
16351		NaN	Living Characters	NaN
16352		NaN	Living Characters	NaN
16353		NaN	Living Characters	NaN
16354		NaN	Deceased Characters	NaN
16355		NaN	Living Characters	NaN
16356		NaN	Living Characters	NaN
16357		NaN	Living Characters	NaN
16358		NaN	Living Characters	NaN
16359		NaN	Living Characters	NaN

16360		NaN	Living Characters	NaN
16361		NaN	Deceased Characters	NaN
16362		NaN	Living Characters	NaN
16363	Homosexual Characters		Living Characters	NaN
16364		NaN	Living Characters	NaN
16365		NaN	Deceased Characters	NaN
16366		NaN	Living Characters	NaN
16367	Homosexual Characters		Living Characters	NaN
16368		NaN	Living Characters	NaN
16369		NaN	Living Characters	NaN
16370		NaN	Living Characters	NaN
16371		NaN	Living Characters	NaN
16372		NaN	Living Characters	NaN
16373		NaN	Living Characters	NaN
16374		NaN	Living Characters	NaN
16375		NaN	Living Characters	NaN

	FIRST APPEARANCE	Year
0	Aug-62	1962.0
1	Mar-41	1941.0
2	Oct-74	1974.0
3	Mar-63	1963.0
4	Nov-50	1950.0
5	Nov-61	1961.0
6	Nov-61	1961.0
7	May-62	1962.0
8	Sep-63	1963.0
9	Nov-61	1961.0
10	Sep-63	1963.0
11	Nov-61	1961.0
12	NaN	NaN
13	May-75	1975.0
14	Sep-64	1964.0
15	Apr-64	1964.0
16	Jul-63	1963.0
17	Jun-65	1965.0
18	Mar-63	1963.0
19	Sep-63	1963.0
20	Jan-62	1962.0
21	Sep-63	1963.0
22	Sep-63	1963.0
23	May-75	1975.0
24	Mar-64	1964.0
25	May-63	1963.0
26	Jun-63	1963.0
27	Sep-63	1963.0
28	Apr-64	1964.0
29	May-75	1975.0

...	...	...
16346	NaN	NaN
16347	NaN	NaN
16348	NaN	NaN
16349	NaN	NaN
16350	NaN	NaN
16351	NaN	NaN
16352	NaN	NaN
16353	NaN	NaN
16354	NaN	NaN
16355	NaN	NaN
16356	NaN	NaN
16357	NaN	NaN
16358	NaN	NaN
16359	NaN	NaN
16360	NaN	NaN
16361	NaN	NaN
16362	NaN	NaN
16363	NaN	NaN
16364	NaN	NaN
16365	NaN	NaN
16366	NaN	NaN
16367	NaN	NaN
16368	NaN	NaN
16369	NaN	NaN
16370	NaN	NaN
16371	NaN	NaN
16372	NaN	NaN
16373	NaN	NaN
16374	NaN	NaN
16375	NaN	NaN

[16376 rows x 13 columns]

In [ ]: