Part-A

December 16, 2019

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

In [2]:
```
# Reading data from input file
data = pd.read_csv("Appointment-No-Show-Data.csv")
print(data.dtypes)
print("----------------------------------------")
data.shape
```

```
PatientId         float64
AppointmentID       int64
Gender             object
ScheduledDay       object
AppointmentDay     object
Age                 int64
Neighbourhood      object
Scholarship         int64
Hipertension        int64
Diabetes            int64
Alcoholism          int64
Handcap             int64
SMS_received        int64
No-show            object
dtype: object
----------------------------------------
```

Out[2]:
```
(110527, 14)
```

In [3]:
```
# Converting the data type to date time
data["ScheduledDay"] = pd.to_datetime(data["ScheduledDay"])
data["AppointmentDay"] = pd.to_datetime(data["AppointmentDay"])
```

In [4]:
```
# Weekday function identifies the day of the week. 0-6 corresponds to Monday-Sunday

data["WeekdayScheduled"] = data["ScheduledDay"].dt.weekday
data["WeekdayAppointment"] = data["AppointmentDay"].dt.weekday
```

```
In [5]: data['isweekdayschedule'] = data['WeekdayScheduled'].apply(lambda x: 1 if x >= 5 else 0

        print('Yes. We have {} entries that are scheduled on a weekend.'.format(sum(data['isweek
        print("-------------------------------------------------------------------------------")
        print("Suggestion : We can delete all those entries for our analysis")
```

```
Yes. We have 24 entries that are scheduled on a weekend.
-------------------------------------------------------------------------------
Suggestion : We can delete all those entries for our analysis
```

```
In [6]: data['isweekdayappointment'] = data['WeekdayAppointment'].apply(lambda x: 1 if x >= 5 e

        print('Yes. We have {} entries that have an appointment on Weekend. These entries are i
        print("-------------------------------------------------------------------------------")
        print("Suggestion : We can delete all those entries for our analysis")
```

```
Yes. We have 39 entries that have an appointment on Weekend. These entries are incorrect.
-------------------------------------------------------------------------------
Suggestion : We can delete all those entries for our analysis
```

```
In [7]: # Part 4
        # (a) - Scheduleday should take place before appointment day
        data["Differencedays"] = data["ScheduledDay"] - data["AppointmentDay"]

        print('Yes. We have {} entries instances where in the appointmentdate is earlier than s
        print("-------------------------------------------------------------------------------")
        print("Suggestion : We can delete all those entries for our analysis")
```

```
Yes. We have 5 entries instances where in the appointmentdate is earlier than scheduled date.
-------------------------------------------------------------------------------
Suggestion : We can delete all those entries for our analysis
```

```
In [8]: # (b) - Gender, handicap and scholarship should have unique value for a patient
        gender_dict = {}
        for i in range(len(data)):
            if data.loc[i, "PatientId"] in gender_dict.keys():
                if(data.loc[i, "Gender"] in gender_dict[data.loc[i, "PatientId"]]):
                    continue
                else:
                    gender_dict[data.loc[i, "PatientId"]].append(data.loc[i, "Gender"])

            else:
                gender_dict[data.loc[i, "PatientId"]] = [data.loc[i,"Gender"]]

        print(sum(1 for i in gender_dict.values() if len(i) >= 2))

        print("Conclusion: None of the patient has difference in gender values multiple visits"
```

```
0
Conclusion: None of the patient has difference in gender values multiple visits
```

```
In [9]: handicap = {}
        for i in range(len(data)):
            if data.loc[i, "PatientId"] in handicap.keys():
                if(data.loc[i, "Handcap"] in handicap[data.loc[i, "PatientId"]]):
                    continue
                else:
                    handicap[data.loc[i, "PatientId"]].append(data.loc[i, "Handcap"])

            else:
                handicap[data.loc[i, "PatientId"]] = [data.loc[i,"Handcap"]]

        print(sum(1 for i in handicap.values() if len(i) >= 2))

        print("Conclusion: None of the patient has difference in handicap values for multiple
```

```
0
Conclusion: None of the patient has difference in handicap values for multiple visits
```

```
In [10]: Scholarship = {}
         for i in range(len(data)):
             if data.loc[i, "PatientId"] in Scholarship.keys():
                 if(data.loc[i, "Scholarship"] in Scholarship[data.loc[i, "PatientId"]]):
                     continue
                 else:
                     Scholarship[data.loc[i, "PatientId"]].append(data.loc[i, "Scholarship"])

             else:
                 Scholarship[data.loc[i, "PatientId"]] = [data.loc[i,"Scholarship"]]

         print(sum(1 for i in Scholarship.values() if len(i) >= 2))

         print("Conclusion: None of the patient has difference in Scholarship values for multi
```

```
0
Conclusion: None of the patient has difference in Scholarship values for multiple visits
```

```
In [11]: age = {}


         for i in range(len(data)):
             if data.loc[i, "PatientId"] in age.keys():
                 #if data.loc[i, "Age"] not in age[data.loc[i, "PatientId"]]:
                     age[data.loc[i, "PatientId"]].append(data.loc[i, "Age"])
```

```
            else:
                age[data.loc[i, "PatientId"]] = [data.loc[i,"Age"]]

        for i in age.keys():
            if(age[i] != sorted(age[i])):
                print(i,age[i])

        # Conclusion : We have entried with mismatch in age.
        # Criteria not met - The age is different in different appointments, then the appoint
```

```
441431798896847.0 [89, 90, 89, 89]
5567159523622.0 [2, 3, 2, 3, 3, 3, 3]
526457271918178.0 [17, 18, 17, 17]
94866145193937.0 [26, 26, 26, 27, 26, 27]
7343477663565.0 [77, 77, 77, 78, 77]
24448321454.0 [1, 1, 1, 2, 2, 1]
91581872616666.0 [34, 35, 34]
81213966782532.0 [47, 47, 47, 47, 48, 48, 48, 48, 48, 48, 48, 48, 48, 47, 48, 48, 47, 48, 48, 4
67793226134888.0 [2, 3, 2, 3]
841147961165.0 [17, 18, 18, 18, 17]
27989988316.0 [38, 38, 39, 38, 38, 38, 38]
9712431915876.0 [21, 22, 21, 22, 22]
726946859782.0 [73, 73, 73, 73, 73, 73, 74, 73, 73, 74, 74]
834798139251.0 [49, 49, 49, 49, 50, 49, 49, 49, 50, 49, 49, 49, 49, 50, 50, 50, 50]
28532485743924.0 [61, 62, 62, 62, 61, 62, 62, 62, 61, 62, 62, 62, 62, 62, 62]
31359178796118.0 [48, 48, 48, 49, 48, 48, 49, 49, 49, 48, 48, 49, 48, 48, 49, 49]
863266615443557.0 [52, 53, 52, 53, 53, 53, 52, 53, 53, 53, 52, 53, 53, 52, 53, 53, 53, 53]
548284689473265.0 [60, 60, 60, 60, 60, 60, 60, 61, 60, 60, 60, 60, 61, 61, 61, 61, 61]
75449195884283.0 [56, 56, 56, 57, 56, 56, 56, 56, 56, 56]
741524498166258.0 [60, 60, 60, 61, 60, 60, 60, 60, 60, 61]
8828534172419.0 [68, 69, 68, 69, 69, 69, 69]
496777386375.0 [56, 56, 57, 57, 56, 57, 57, 57, 56, 57, 57, 57, 57, 57, 57, 57, 57, 57, 57, 57
258424392677.0 [49, 49, 49, 49, 50, 50, 50, 50, 49, 50, 50, 49, 50, 49, 49, 50, 49, 50, 50, 49
871374938638855.0 [63, 63, 63, 63, 64, 63, 63, 63, 64, 63, 63, 63, 63, 63, 63, 63, 63, 63,
73962269658754.0 [41, 42, 41, 42]
554333725686493.0 [52, 51]
1764194433642.0 [20, 19]
442654573782243.0 [59, 58, 58, 59]
673364656397.0 [28, 28, 28, 28, 27]
284187334839.0 [36, 35, 35]
3176755233235.0 [88, 88, 88, 87, 87]
262638886167187.0 [19, 18]
596853753814.0 [35, 35, 35, 34]
136182642528387.0 [49, 48, 49]
13487876248348.0 [51, 50]
2459844361911.0 [66, 65, 65]
99427113877488.0 [57, 57, 58, 57]
```

3571437593233.0 [70, 69, 70]
68391591311966.0 [47, 46, 47, 46, 47, 47]
826482348319264.0 [31, 32, 31, 31, 32, 32]
324872673962117.0 [53, 52, 52, 52, 53, 53, 52, 52, 53, 52, 52, 53, 53, 53]
78775287514972.0 [27, 27, 26, 27]
986162815579582.0 [45, 45, 44, 45, 44, 44, 45, 44, 45, 45, 45, 45, 44, 45, 45, 45, 44, 45, 44,
72371443479854.0 [40, 39, 39, 40, 40]
945287579358673.0 [39, 40, 39]
13264579795931.0 [74, 74, 73, 74, 73, 74, 74]
28695244691676.0 [67, 66, 67]
2246258522789.0 [55, 54, 55, 55]
6143344679435.0 [82, 81, 82]
51775247878421.0 [54, 53, 54, 54]
95194928669972.0 [70, 69, 69, 69, 69, 70]
247852377364167.0 [64, 65, 64, 64, 65]
491989729831692.0 [59, 58, 58, 59, 59]
91837389522927.0 [49, 50, 49, 50, 49]
37412972916427.0 [78, 77]
1498811933864.0 [50, 49, 49, 49, 49, 49, 50, 50, 50, 50]
412145433477.0 [36, 37, 36, 37, 37]
326648852377965.0 [24, 23, 24, 24, 24]
3965322382843.0 [8, 7]
56211292273458.0 [8, 9, 8, 9]
3257918658792.0 [8, 8, 7]
51841113266722.0 [8, 7]
54443767614768.0 [43, 42]
29135353999775.0 [22, 23, 22, 23, 23]
77619232722763.0 [28, 28, 29, 28, 28, 29]
8852723712332.0 [33, 32]
551725965595814.0 [63, 62]
4355693229845.0 [9, 9, 10, 9, 10, 10]
3345256473.0 [39, 38]
37994336685892.0 [61, 61, 61, 61, 62, 61, 61, 61, 61, 61, 61, 61, 61, 61, 61, 61, 61, 62, 62, 6
522596789798239.0 [39, 38]
37943534192721.0 [9, 8]
9133657377593.0 [52, 52, 53, 52, 53]
85784143433277.0 [31, 31, 31, 31, 31, 31, 32, 32, 31, 32, 32, 31, 31, 32]
26615199966515.0 [33, 32, 33]
3741693143169.0 [55, 56, 56, 56, 55, 56, 56]
72851223342246.0 [55, 56, 55, 56]
9326397647852.0 [27, 27, 27, 27, 27, 27, 28, 27, 27, 27, 27, 27, 28, 28, 28, 28, 28]
144179259631412.0 [45, 45, 44, 45]
3554216592229.0 [34, 34, 33, 33, 33]
461724434639.0 [2, 1, 1, 1]
814716826978881.0 [10, 11, 10]
555364118193729.0 [17, 16, 16]
4352739845586.0 [43, 42]
74446587851354.0 [33, 32, 32, 33, 33]

5239216287933.0 [80, 79, 80]
952788794771451.0 [79, 80, 79]
3615142518775.0 [23, 24, 23, 24, 24]
623893427438.0 [25, 25, 24, 24]
215844296214149.0 [37, 36, 36]
348771362758491.0 [51, 50, 51, 51]
148331859466744.0 [54, 53, 53, 54]
23895254172.0 [59, 58]
5765346781124.0 [1, 0, 1]
171936517145444.0 [67, 66, 67]
9525694478154.0 [70, 69]
422812612546.0 [56, 57, 57, 56]
4821115864239.0 [7, 7, 8, 7, 8]
2977941938987.0 [43, 43, 42]
23519815538343.0 [1, 0, 1]
636735389134779.0 [76, 76, 76, 76, 75]
9629299847949.0 [40, 39]
686927171986418.0 [30, 29]
11238367556569.0 [29, 28, 28]
673788649699968.0 [29, 28]
795491396711336.0 [26, 26, 26, 25, 25]
1859211848923.0 [38, 37, 37, 37, 38]
647668892475693.0 [46, 46, 45, 45]
785592131774334.0 [31, 31, 30, 30, 31, 31]
94411894549983.0 [87, 87, 88, 87, 88, 88, 88, 88]
46687682547277.0 [31, 31, 31, 30]
3526815486217.0 [49, 49, 48, 49]
96131746626662.0 [38, 38, 38, 39, 39, 39, 38, 39, 39]
48249386169945.0 [53, 52]
2213643753729.0 [44, 43, 44]
8714199656817.0 [53, 52, 52, 53]
7783718946167.0 [19, 18]
94456576941976.0 [5, 4]
319479322852564.0 [37, 37, 36, 37, 37]
914477235417273.0 [35, 34]
49976884839224.0 [18, 17, 18]
99575991136368.0 [44, 43]
71185353265575.0 [38, 37, 37]
3592971226558.0 [44, 43]
9888222214976.0 [1, 1, 1, 2, 1]
517567676968.0 [17, 17, 17, 17, 18, 17, 17, 17, 17, 17, 18, 18]
2183482638962.0 [60, 60, 59, 60]
53534197453837.0 [50, 50, 51, 51, 50]
86169866633.0 [86, 87, 87, 86, 87]
199429275315785.0 [45, 44, 45]
27138244415357.0 [49, 50, 49]
1217364296.0 [49, 48, 49, 49]
85484221951869.0 [30, 30, 29, 30]

71724519836714.0 [44, 44, 44, 44, 45, 44, 45, 45, 45]
1835776214631.0 [72, 72, 72, 72, 71, 72, 72]
67662267316191.0 [25, 25, 26, 25]
71842123437414.0 [45, 45, 45, 44, 45]
56135513537833.0 [26, 27, 27, 27, 26, 27]
53151829685666.0 [60, 59, 60, 60]
25488622482828.0 [24, 24, 23, 23]
429998399768479.0 [17, 16]
97173476551.0 [14, 15, 15, 14, 15, 15]
2435946356617.0 [42, 43, 42, 42, 42, 42, 42, 42, 42, 43, 43, 43]
876979717673738.0 [22, 22, 23, 22, 23, 23]
22878144792827.0 [15, 15, 15, 15, 15, 15, 16, 15]
16798678591586.0 [73, 72, 73, 73, 73, 73, 73, 73, 73]
79219263716974.0 [59, 58, 59]
72589141991386.0 [9, 10, 9, 10]
191874988484646.0 [33, 33, 32, 32, 33]
868126583293.0 [30, 30, 31, 30]
799771532794835.0 [4, 5, 4, 5]
357552918572233.0 [46, 47, 46, 46, 46]
2797782164645.0 [25, 26, 25]
385715913296212.0 [50, 51, 50]
954464386155548.0 [48, 49, 49, 48, 49]
772134969567489.0 [24, 23]
62636562322246.0 [11, 11, 10]
8231719732637.0 [42, 43, 43, 42]
4761525653111.0 [14, 13]
886162945474278.0 [8, 7]
2458268636912.0 [53, 53, 54, 53, 53, 53, 54]
45872591783433.0 [42, 42, 42, 42, 42, 42, 43, 42, 43, 43, 43]
339697273859477.0 [58, 57, 58]
153522341971323.0 [28, 29, 29, 28]
7233373911188.0 [28, 27, 28, 28]
539542451692.0 [42, 42, 42, 42, 42, 41, 42]
8273238259516.0 [3, 2, 3]
154368918871336.0 [38, 37]
91129344118.0 [24, 23]
8922632122719.0 [4, 3]
33488149875846.0 [82, 81]
52413972636433.0 [52, 51]
3139597142774.0 [51, 50, 51]
923768744371225.0 [72, 71, 72]
534928747131878.0 [80, 79, 79, 79, 80, 80, 80, 80]
425917412221.0 [51, 50, 51, 50, 51, 51, 50, 50]
93158632513652.0 [21, 22, 21]
58318454478879.0 [39, 38]
7218811271376.0 [81, 80, 81, 81]
78796728799266.0 [53, 52, 53]
8468613288842.0 [22, 22, 23, 22, 22, 22, 22, 22, 23, 23, 23, 23]

1227457867912.0 [43, 43, 43, 42, 43, 43, 43, 43, 43, 43]
59323741249917.0 [65, 65, 65, 65, 65, 66, 65, 65, 65, 65, 65, 65, 65, 66, 66, 66, 66]
1977936998781.0 [61, 62, 61, 61, 61, 62, 62]
86458552841654.0 [27, 27, 27, 28, 27, 27, 28, 27, 28, 28, 28]
933599918554.0 [59, 58]
64345423327588.0 [51, 51, 50]
88883712788189.0 [90, 89]
5197889628921.0 [1, 0]
12915328916148.0 [74, 73]
84653719847945.0 [86, 85, 86, 86, 86]
85258359996369.0 [71, 70, 70, 71, 71]
536165438892.0 [18, 17]
95332789523636.0 [29, 29, 29, 30, 29, 30, 30, 30]
6849728138464.0 [7, 6, 6]
39877958212979.0 [75, 74]
916623214957.0 [42, 42, 41, 42, 42, 42, 42]
462516862932765.0 [15, 16, 15, 16, 16, 16, 16]
779528544111143.0 [37, 37, 36, 36]
48556922666934.0 [45, 44, 44, 44, 45, 45, 45, 45]
9851772915268.0 [23, 23, 22]
41217995545889.0 [9, 8]
5454976421989.0 [28, 27, 28]
224493451777.0 [35, 36, 35, 35, 36, 35]
43327939247761.0 [58, 58, 58, 59, 59, 58, 58, 58, 59, 59]
427581952897615.0 [65, 66, 66, 66, 66, 65, 66, 66, 66, 66, 66]
1126541547466.0 [54, 54, 53]
583862231194528.0 [17, 17, 17, 18, 18, 17]
457352921116.0 [18, 17]
848888839915289.0 [40, 39]
178725979971382.0 [83, 82]
49111362868275.0 [52, 51]
314453617858619.0 [17, 17, 16, 16]
69483178233285.0 [27, 26]
833227727921254.0 [1, 1, 0, 1, 1, 1]
1998677482451.0 [62, 61]
87751781373969.0 [58, 57, 58, 58, 58]
87841434935852.0 [2, 1, 2]
95658423471799.0 [68, 67]
94759232137264.0 [51, 50, 51]
23864951634163.0 [4, 3]
86138665742284.0 [64, 65, 64]
46455714367665.0 [15, 15, 14, 15]
875359764898749.0 [9, 8, 9]
58881255967391.0 [44, 43, 43, 44, 44, 44]
539139981641122.0 [3, 2, 2, 3, 3]
6675118595731.0 [38, 37]
898589715173887.0 [19, 18]
25387937196542.0 [5, 4]

28531287317177.0 [1, 0, 1]
23758911576891.0 [1, 0]
116959988999174.0 [40, 40, 41, 40]
3515954487176.0 [29, 30, 29]
257693717845.0 [53, 52, 53]
593897591752.0 [22, 22, 21]
34867754122127.0 [41, 40]
88595623726224.0 [35, 35, 34, 35]
2277762598537.0 [7, 6]
274348768898463.0 [16, 15]
7374275369434.0 [17, 16, 17]
911594989391.0 [44, 43, 43]
56522486875269.0 [76, 75, 76]
238996137394148.0 [28, 28, 27, 28]
793212543441396.0 [26, 25]
68545173979.0 [17, 18, 17]
43199387596339.0 [63, 62]
49194942193312.0 [35, 34]
4214836353162.0 [65, 64]
3899959422895.0 [43, 42]
16372358139357.0 [0, 1, 0, 0, 0]
24661431761226.0 [60, 59]
789515756169569.0 [68, 67, 67]
2796337397542.0 [58, 57, 58, 58, 58, 58]
465132113185454.0 [13, 13, 12]
54281515617532.0 [36, 37, 36]
36192893592111.0 [29, 30, 29, 30]
614343448988718.0 [6, 7, 6, 7]
29127973566646.0 [76, 75, 76, 75]
392739499447919.0 [54, 53]
22831479979362.0 [71, 71, 72, 71, 72]
541829916589.0 [36, 36, 37, 36]
9884913194.0 [6, 6, 5, 6, 6]
74587547259358.0 [34, 33]
16716526114164.0 [30, 29]
16329323133156.0 [20, 19, 20]
6446622277797.0 [54, 54, 55, 54]
197566864898.0 [32, 32, 32, 32, 32, 32, 33, 32, 32, 33]
685996785123.0 [58, 57]
897641496984939.0 [3, 2]
13118936499174.0 [5, 4, 5]
9442336472761.0 [47, 46, 47]
725983755516.0 [19, 19, 19, 19, 18, 19, 19, 19, 18, 19, 18, 19, 18, 19, 19, 19, 19, 19, 19, 19
72614136787532.0 [30, 29, 30]
148418344786.0 [6, 5]
646197667728973.0 [16, 15]
283113846355781.0 [25, 25, 25, 26, 26, 25, 26, 26]
89541443642743.0 [81, 80]

228553541646274.0 [60, 61, 60]
49857225882957.0 [76, 75, 76, 75, 76, 76, 76, 76, 76]
239552727888.0 [74, 73, 73, 73, 73, 73, 73, 74]
5342896776979.0 [65, 66, 65, 65, 66, 66, 66]
1666518766336.0 [77, 77, 78, 77, 77, 78, 78, 77, 77, 78, 78]
779478897631643.0 [63, 63, 62]
121174219573222.0 [82, 81]
678391415599.0 [35, 34, 34]
647114427237477.0 [27, 26, 26]
51114476492581.0 [50, 49, 49]
425613399554188.0 [4, 4, 3]
276587384841131.0 [50, 49, 50, 49, 49, 50]
16936715442791.0 [45, 44, 44, 45]
9381876758637.0 [10, 9, 9, 10, 10, 10]
823681776976369.0 [13, 14, 13, 14, 14, 14]
591733659913661.0 [0, 1, 0, 1, 1]
66815697866759.0 [12, 12, 11]
497276854484.0 [15, 16, 16, 15]
31598816735479.0 [34, 33]
31791925616328.0 [57, 56]
999479168794227.0 [44, 45, 44]
52599251296133.0 [56, 55, 56, 56, 56]
944226174674.0 [24, 23, 24]
1643668346244.0 [78, 77]
1198157171975.0 [10, 10, 10, 10, 10, 10, 11, 10, 10, 10, 10, 10, 11, 10, 10, 11, 11, 11, 11, 1
1823656361699.0 [11, 11, 11, 10, 11, 11]
53696749534376.0 [9, 8, 9, 9]
84947544118237.0 [7, 8, 7, 7, 8, 7, 7, 8, 7, 7, 8]
59573646256238.0 [5, 4, 5, 5, 4, 5]
562268283176744.0 [17, 16, 16, 17, 17, 17, 17]
552121474263.0 [60, 59]
314527232856552.0 [10, 9]
842294755532315.0 [30, 29, 30, 30]
39961733268143.0 [16, 15]
7487854375951.0 [18, 17]
81428455459752.0 [59, 58, 58, 59]
41531353974974.0 [5, 4, 5]
97137488744159.0 [24, 23, 23]
75612146234696.0 [8, 7, 8]
371545547125.0 [32, 31, 32]
329517679422892.0 [21, 21, 20, 21, 21]
66178265367648.0 [44, 44, 43]
894338462663844.0 [23, 22, 22, 23]
36187286318946.0 [52, 51, 52]
819269171894553.0 [17, 16]
89797152627.0 [23, 23, 22]
34111211257574.0 [28, 27, 27]
64256657641911.0 [56, 57, 56, 56]

```
533748128563971.0 [36, 35, 35, 36, 35]
512126575678662.0 [42, 41, 42, 42, 41, 42, 42]
59416452434634.0 [47, 46]
769731576114186.0 [65, 65, 65, 65, 64, 64, 65, 65]
42378651295177.0 [12, 11, 11, 12, 11, 11, 12, 12]
5836366357913.0 [72, 72, 71]
63791599117.0 [33, 33, 33, 34, 33]
2687981344286.0 [60, 59]
47228863177.0 [51, 51, 50, 51, 51, 51, 51, 51]
58546723645496.0 [6, 5]
881431216978.0 [1, 0, 1, 1]
1286321267.0 [34, 33]
81847674321826.0 [17, 16]
1313251118148.0 [47, 46, 47]
88513212492984.0 [53, 52, 53]
96128338755471.0 [36, 36, 35]
4117439975384.0 [80, 79]
445419621221545.0 [2, 1]
46987719578558.0 [1, 0, 0]
83266638139533.0 [9, 8]
15385746248142.0 [68, 69, 69, 68, 68, 69, 69]
33292792573954.0 [39, 38]
2344348428464.0 [71, 70, 71]
777478245323485.0 [44, 43, 43, 43]
38145463433411.0 [61, 60, 61]
82549633651698.0 [0, 1, 0]
422633928124753.0 [51, 51, 50]
82612694333264.0 [51, 50, 51]
837881657858.0 [47, 48, 47]
312597853269988.0 [57, 56, 56, 56, 57, 57]
692757522244.0 [20, 19, 19, 19, 19]
588177487514.0 [38, 37, 37]
293318845296.0 [12, 11]
1611158895.0 [7, 8, 7]
12423479616262.0 [5, 4]
577498242433.0 [11, 10]
967821826529337.0 [2, 1]
73657497536994.0 [36, 35]
9286653679.0 [37, 36]
778564276439829.0 [1, 0]
7765845965853.0 [30, 30, 29]
4397165563425.0 [66, 66, 67, 66]
993256795751.0 [24, 23]
978523623546.0 [49, 50, 49]
4964336865591.0 [7, 6]
```

```
In [12]: print("Conclusion: We have entried with mismatch in age. Criteria not met - The age is
```

Conclusion: We have entried with mismatch in age. Criteria not met - The age is different in di

In [13]:
```
# Part 5
# Negative ages
data[data["Age"] < 0]

print('Yes. We have {} entries where patient is having negative age.'.format(len(data
print("--------------------------------------------------------------------------")
print("Suggestion : We can delete all those entries for our analysis")
```

```
Yes. We have 1 entries where patient is having negative age.
--------------------------------------------------------------------------
Suggestion : We can delete all those entries for our analysis
```

In [14]:
```
# Part 6 - Removing people with age greater than 100 as they are outliers
print('Yes. We have {} entries where patient is having 100+ age.'.format(len(data[data
print("Removing people with age greater than 100")
data = data[data["Age"] <100]
print(data.shape)
```

```
Yes. We have 7 entries where patient is having 100+ age.
Removing people with age greater than 100
(110516, 19)
```