

Join GitHub today

Dismiss

GitHub is home to over 28 million developers working together to host and review code, manage projects, and build software together.

Sign up

Branch: master ▾

solutions / Mission218Solution.ipynb

Find file

Copy path

 **srinify** small solutions fix

f021395 on Mar 4, 2018

1 contributor

476 lines (475 sloc) 12.6 KB

US Gun Deaths Guided Project Solutions

Introducing US Gun Deaths Data

```
In [30]: import csv

with open("guns.csv", "r") as f:
    reader = csv.reader(f)
    data = list(reader)
```

```
In [31]: print(data[:5])

[['', 'year', 'month', 'intent', 'police', 'sex', 'age', 'race', 'hispanic', 'place', 'educatio
n'], ['1', '2012', '01', 'Suicide', '0', 'M', '34', 'Asian/Pacific Islander', '100', 'Home', '4'],
['2', '2012', '01', 'Suicide', '0', 'F', '21', 'White', '100', 'Street', '3'], ['3', '2012', '01',
'Suicide', '0', 'M', '60', 'White', '100', 'Other specified', '4'], ['4', '2012', '02', 'Suicide',
'0', 'M', '64', 'White', '100', 'Home', '4']]
```

Removing Headers From A List Of Lists

```
In [32]: headers = data[:1]
data = data[1:]
print(headers)
print(data[:5])

[['', 'year', 'month', 'intent', 'police', 'sex', 'age', 'race', 'hispanic', 'place', 'educatio
n']]
[['1', '2012', '01', 'Suicide', '0', 'M', '34', 'Asian/Pacific Islander', '100', 'Home', '4'],
['2', '2012', '01', 'Suicide', '0', 'F', '21', 'White', '100', 'Street', '3'], ['3', '2012', '01',
'Suicide', '0', 'M', '60', 'White', '100', 'Other specified', '4'], ['4', '2012', '02', 'Suicide',
'0', 'M', '64', 'White', '100', 'Home', '4'], ['5', '2012', '02', 'Suicide', '0', 'M', '31', 'Whit
e', '100', 'Other specified', '2']]
```

Counting Gun Deaths By Year

```
In [33]: years = [row[1] for row in data]

year_counts = {}
for year in years:
    if year not in year_counts:
        year_counts[year] = 1
    else:
        year_counts[year] += 1

year_counts
```

```
Out[33]: {'2012': 33563, '2013': 33636, '2014': 33599}
```

Exploring Gun Deaths By Month And Year

```
In [34]: import datetime

dates = [datetime.datetime(year=int(row[1]), month=int(row[2]), day=1) for row in data]
dates[:5]
```

```
Out[34]: [datetime.datetime(2012, 1, 1, 0, 0),
datetime.datetime(2012, 1, 1, 0, 0),
datetime.datetime(2012, 1, 1, 0, 0),
datetime.datetime(2012, 2, 1, 0, 0),
datetime.datetime(2012, 2, 1, 0, 0)]
```

```
In [35]: date_counts = {}

for date in dates:
    if date not in date_counts:
        date_counts[date] = 0
    date_counts[date] += 1

date_counts
```

```
Out[35]: {datetime.datetime(2012, 1, 1, 0, 0): 2758,
datetime.datetime(2012, 2, 1, 0, 0): 2357,
datetime.datetime(2012, 3, 1, 0, 0): 2743,
datetime.datetime(2012, 4, 1, 0, 0): 2795,
```

```

datetime.datetime(2012, 5, 1, 0, 0): 2999,
datetime.datetime(2012, 6, 1, 0, 0): 2826,
datetime.datetime(2012, 7, 1, 0, 0): 3026,
datetime.datetime(2012, 8, 1, 0, 0): 2954,
datetime.datetime(2012, 9, 1, 0, 0): 2852,
datetime.datetime(2012, 10, 1, 0, 0): 2733,
datetime.datetime(2012, 11, 1, 0, 0): 2729,
datetime.datetime(2012, 12, 1, 0, 0): 2791,
datetime.datetime(2013, 1, 1, 0, 0): 2864,
datetime.datetime(2013, 2, 1, 0, 0): 2375,
datetime.datetime(2013, 3, 1, 0, 0): 2862,
datetime.datetime(2013, 4, 1, 0, 0): 2798,
datetime.datetime(2013, 5, 1, 0, 0): 2806,
datetime.datetime(2013, 6, 1, 0, 0): 2920,
datetime.datetime(2013, 7, 1, 0, 0): 3079,
datetime.datetime(2013, 8, 1, 0, 0): 2859,
datetime.datetime(2013, 9, 1, 0, 0): 2742,
datetime.datetime(2013, 10, 1, 0, 0): 2808,
datetime.datetime(2013, 11, 1, 0, 0): 2758,
datetime.datetime(2013, 12, 1, 0, 0): 2765,
datetime.datetime(2014, 1, 1, 0, 0): 2651,
datetime.datetime(2014, 2, 1, 0, 0): 2361,
datetime.datetime(2014, 3, 1, 0, 0): 2684,
datetime.datetime(2014, 4, 1, 0, 0): 2862,
datetime.datetime(2014, 5, 1, 0, 0): 2864,
datetime.datetime(2014, 6, 1, 0, 0): 2931,
datetime.datetime(2014, 7, 1, 0, 0): 2884,
datetime.datetime(2014, 8, 1, 0, 0): 2970,
datetime.datetime(2014, 9, 1, 0, 0): 2914,
datetime.datetime(2014, 10, 1, 0, 0): 2865,
datetime.datetime(2014, 11, 1, 0, 0): 2756,
datetime.datetime(2014, 12, 1, 0, 0): 2857}

```

Exploring Gun Deaths By Race And Sex

```

In [54]: sexes = [row[5] for row in data]
sex_counts = {}
for sex in sexes:
    if sex not in sex_counts:
        sex_counts[sex] = 0
    sex_counts[sex] += 1
sex_counts

```

```

Out[54]: {'F': 14449, 'M': 86349}

```

```

In [36]: races = [row[7] for row in data]
race_counts = {}
for race in races:
    if race not in race_counts:
        race_counts[race] = 0
    race_counts[race] += 1
race_counts

```

```

Out[36]: {'Asian/Pacific Islander': 1326,
'Black': 23296,
'Hispanic': 9022,
'Native American/Native Alaskan': 917,
'White': 66237}

```

Findings so far

Gun deaths in the US seem to disproportionately affect men vs women. They also seem to disproportionately affect minorities, although having some data on the percentage of each race in the overall US population would help.

There appears to be a minor seasonal correlation, with gun deaths peaking in the summer and declining in the winter. It might be useful to filter by intent, to see if different categories of intent have different correlations with season, race, or gender.

Reading In A Second Dataset

```

In [57]: import csv

with open("census.csv", "r") as f:
    reader = csv.reader(f)
    census = list(reader)

census

```

```

Out[57]: [['Id',
'Year',
'...'

```

```

'Id',
'Sex',
'Id',
'Hispanic Origin',
'Id',
'Id2',
'Geography',
'Total',
'Race Alone - White',
'Race Alone - Hispanic',
'Race Alone - Black or African American',
'Race Alone - American Indian and Alaska Native',
'Race Alone - Asian',
'Race Alone - Native Hawaiian and Other Pacific Islander',
'Two or More Races'],
['cen42010',
'April 1, 2010 Census',
'totsex',
'Both Sexes',
'tothisp',
'Total',
'0100000US',
'',
'United States',
'308745538',
'197318956',
'44618105',
'40250635',
'3739506',
'15159516',
'674625',
'6984195']]

```

Computing Rates Of Gun Deaths Per Race

```

In [40]: mapping = {
    "Asian/Pacific Islander": 15159516 + 674625,
    "Native American/Native Alaskan": 3739506,
    "Black": 40250635,
    "Hispanic": 44618105,
    "White": 197318956
}

race_per_hundredk = {}
for k,v in race_counts.items():
    race_per_hundredk[k] = (v / mapping[k]) * 100000

race_per_hundredk

```

```

Out[40]: {'Asian/Pacific Islander': 8.374309664161762,
'Black': 57.8773477735196,
'Hispanic': 20.220491210910907,
'Native American/Native Alaskan': 24.521955573811088,
'White': 33.56849303419181}

```

Filtering By Intent

```

In [41]: intents = [row[3] for row in data]
homicide_race_counts = {}
for i,race in enumerate(races):
    if race not in homicide_race_counts:
        homicide_race_counts[race] = 0
    if intents[i] == "Homicide":
        homicide_race_counts[race] += 1

race_per_hundredk = {}
for k,v in homicide_race_counts.items():
    race_per_hundredk[k] = (v / mapping[k]) * 100000

race_per_hundredk

```

```

Out[41]: {'Asian/Pacific Islander': 3.530346230970155,
'Black': 48.471284987180944,
'Hispanic': 12.627161104219914,
'Native American/Native Alaskan': 8.717729026240365,
'White': 4.6356417981453335}

```

Findings

It appears that gun related homicides in the US disproportionately affect people in the Black and Hispanic racial categories.

