

EDA ON SUICIDES IN INDIA

Presented by: Janani Jennifer





INTRODUCTION

SYSTEM DESIGN

IMPLEMENTATION

CONCLUSION

AGENDA



INTRODUCTION

N

BUSINESS UNDERSTANDING

DATA

COLLECTION
DATA

CLEANING
EXPLORATORY DATA ANALYSIS

(EDA)
DATA VISUALIZATION

OUTCOMES AND CONCLUSION

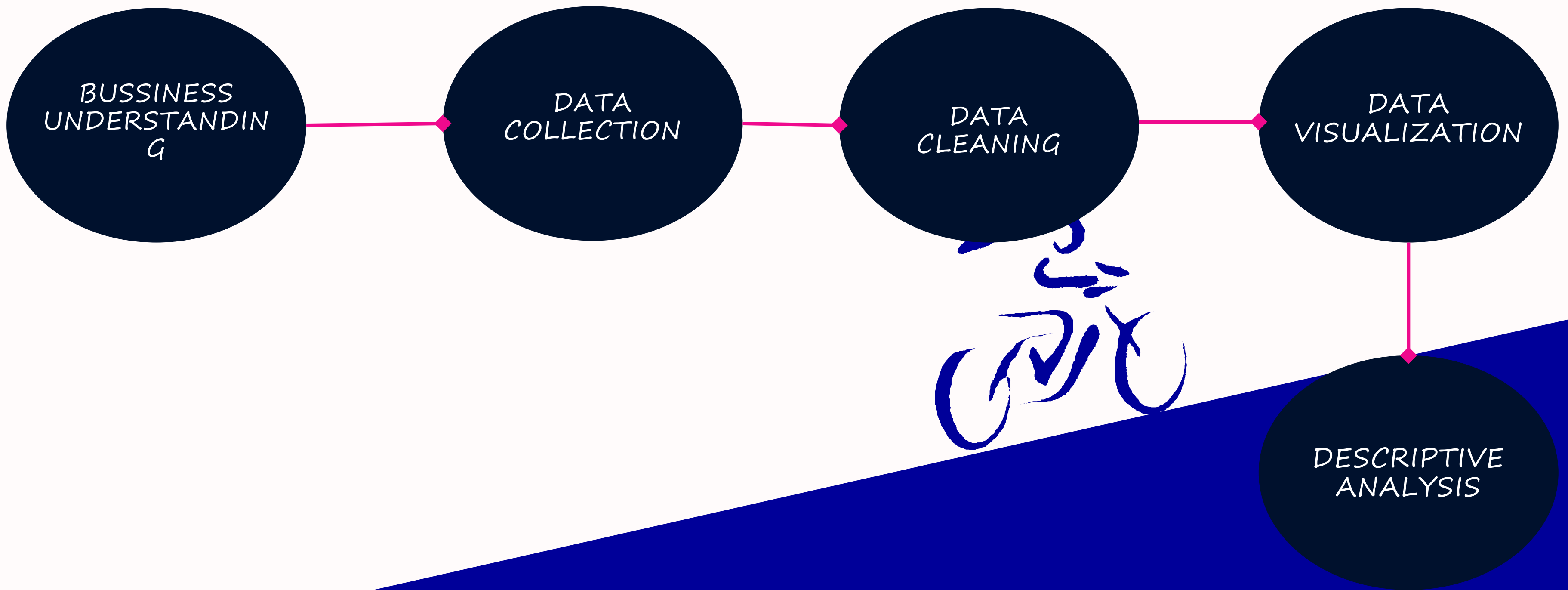
EDA ON SUICIDES



The EDA consists of 7 columns. Here we have 7 features using this we start data analysis of SUICIDES in India. This data set consists of information.

- STATE
- YEAR
- TYPE_CODE
- TYPE
- GENDER
- AGE_GROUP
- TOTAL

SYSTEM DESIGN



IMPLEMENTATION

AIM :

To understand the influence of various factors that would end up taking wrong decision

REQUIRED LIBRARIES TO BE
IMPORTED :

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```



```
df=pd.read_csv("G:\datascience\Suicides.csv
```

READING THE DATA FROM CSV FILE :

(237519, 7)

SIZE IF THE DATA :

LET US CHECK FOR ANY MISSING VALUES

#code

```
df.isna().sum  
()
```

AS SEEN , BELOW THERE ARE NO MISSING (NULL)
VALUES IN THIS DATAFRAME

```
STATE      0  
YEAR      0  
TYPE_CODE  0  
TYPE      0  
GENDER     0  
AGE_GROUP  0  
TOTAL     0
```

```
\DTYPE: INT64
```

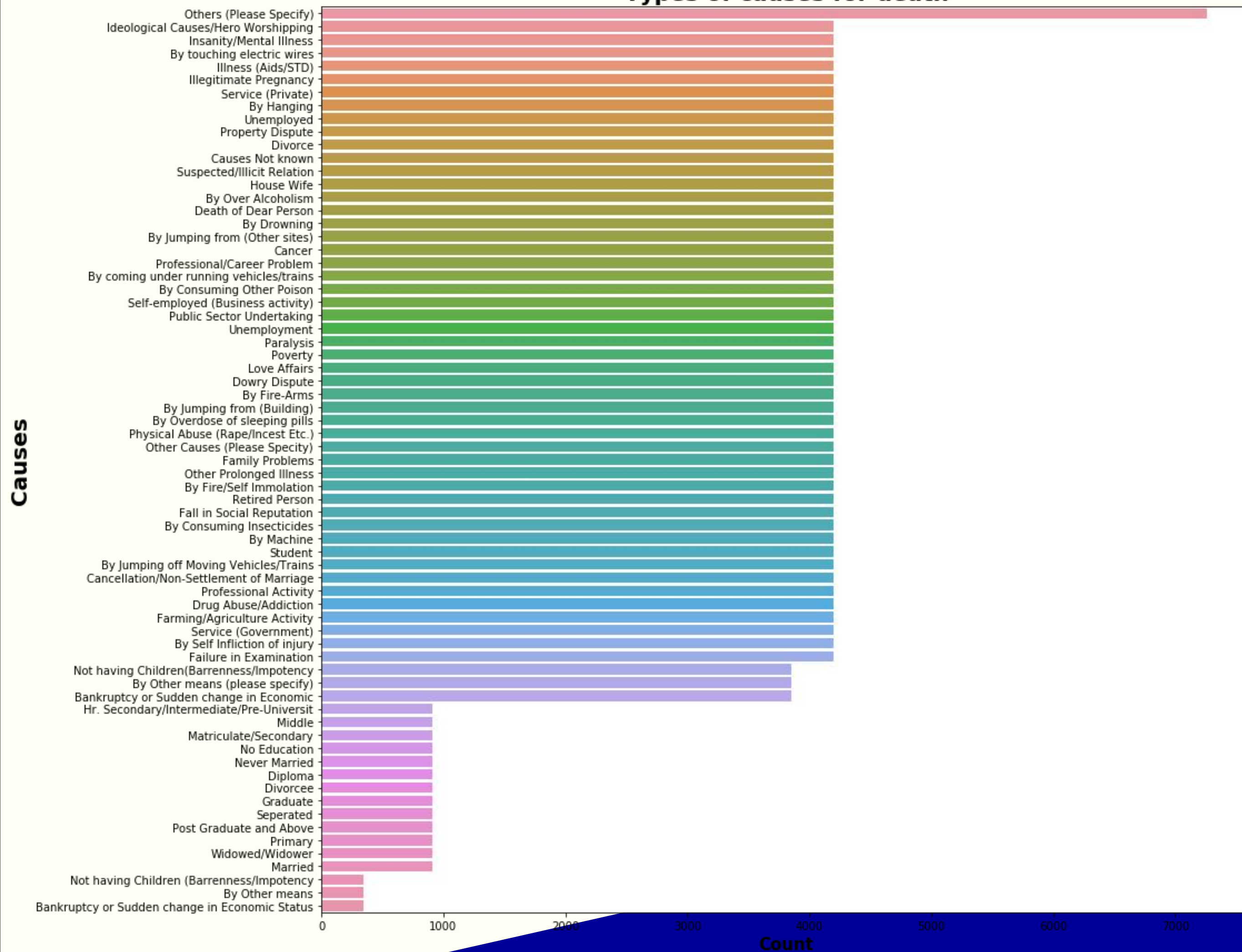

EXPLORATORY DATA ANALYSIS



Q1.

Let us check the
maximum causes of
death

Types of causes for death



- As observed above we can say that the maximum death is caused by 'OTHERS'
- The minimum death is caused by three means of disase

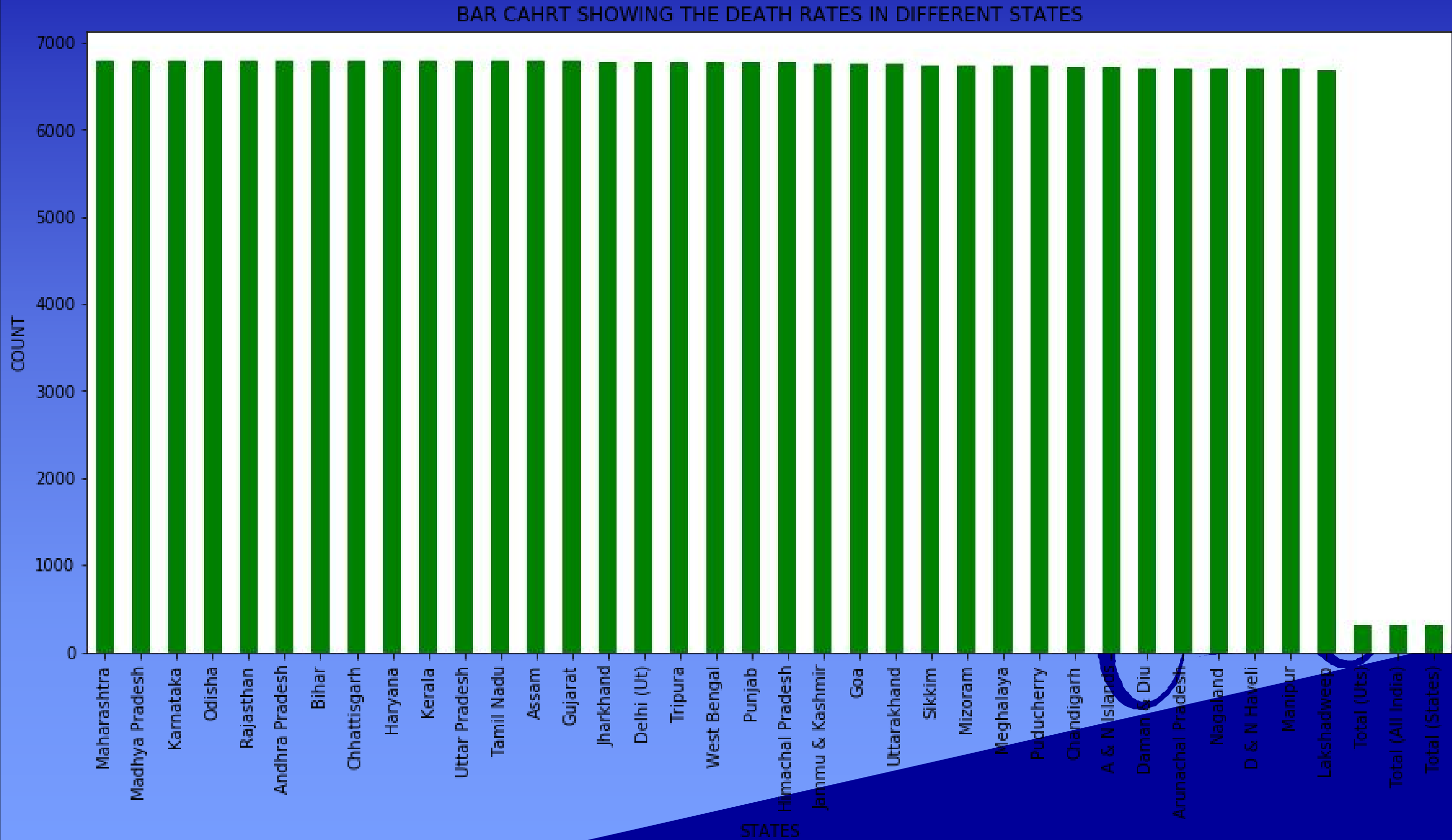
Q2.

Let us know which state has the maximum death rates



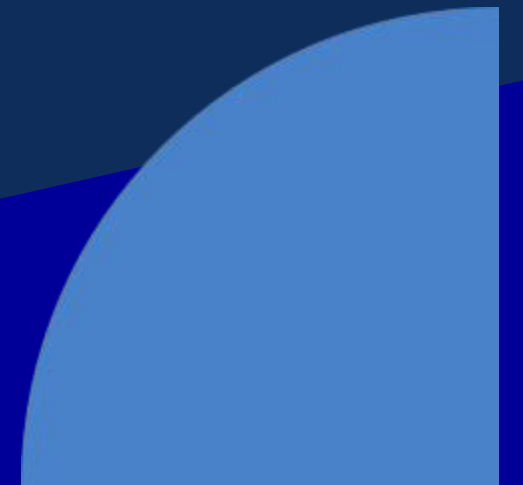
As observed above we can say that the maximum death is in 3 States which is

- Maharashtra (6792)
- Madhya Pradesh (6792)
- Karnataka (6792)



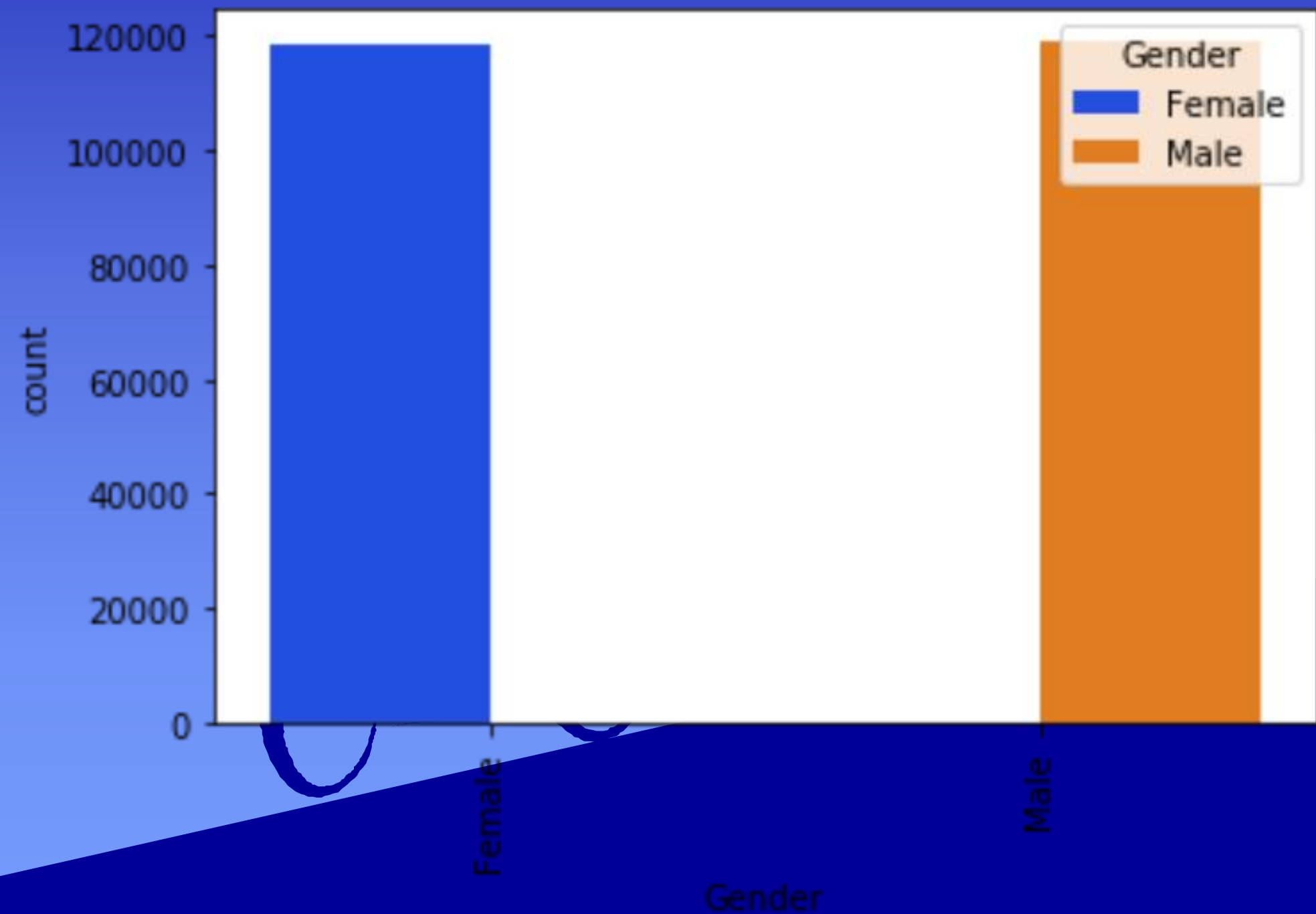
Q3.

Let us see the highest death rates among Male and Female



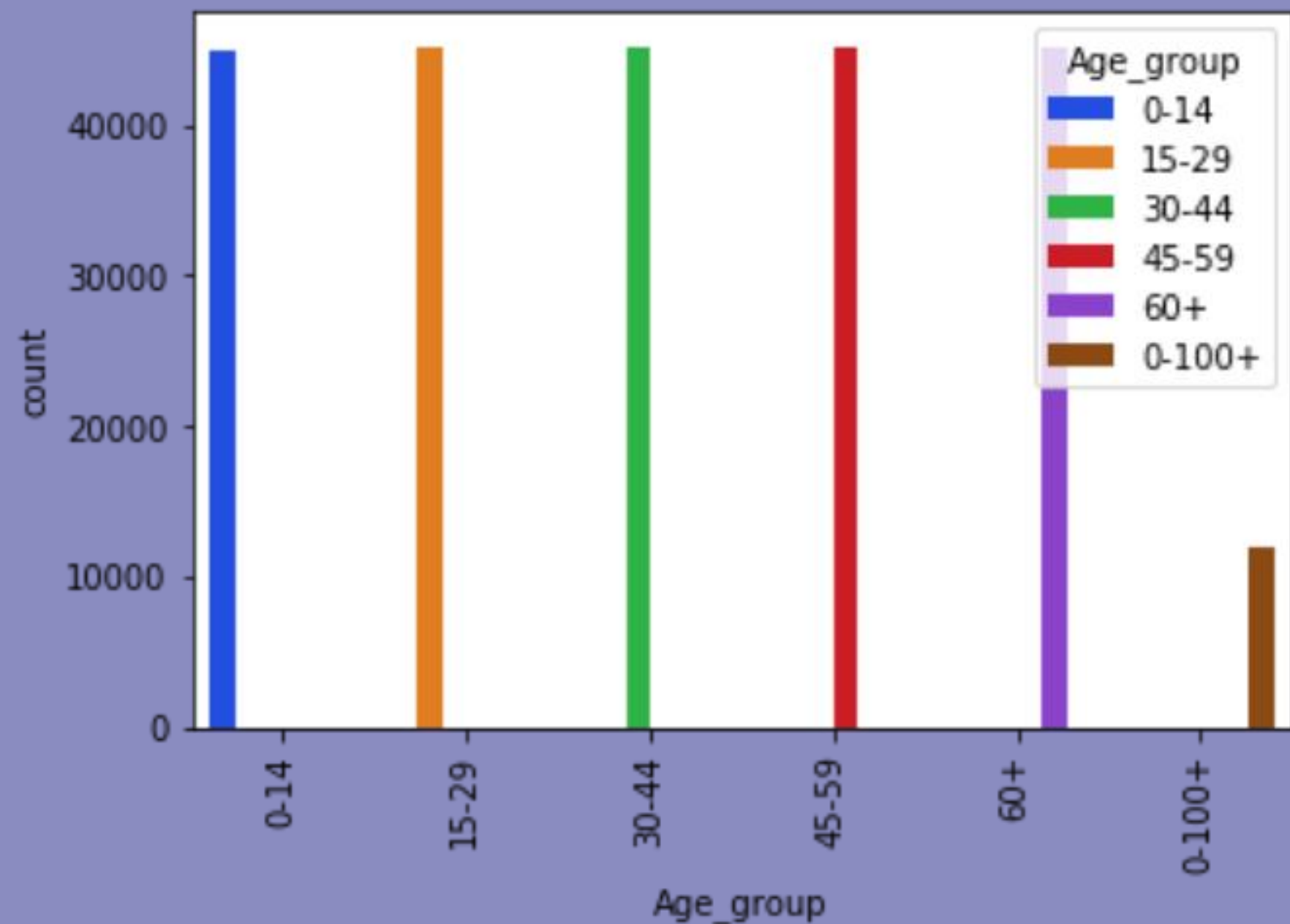
The maximum
deaths among Male
and Female is Male
with rate of 118879

```
Male    118879  
Female  118640  
Name: Gender, dtype:  
int64
```



Q4. Maximum Death rates among different ages





Maximum Death rates is among the age group of 15-29

```
15-29    45223
30-44    45193
45-59    45146
60+      45074
0-14     45027
0-100+   11856
```

Name: Age_group, dtype: int64

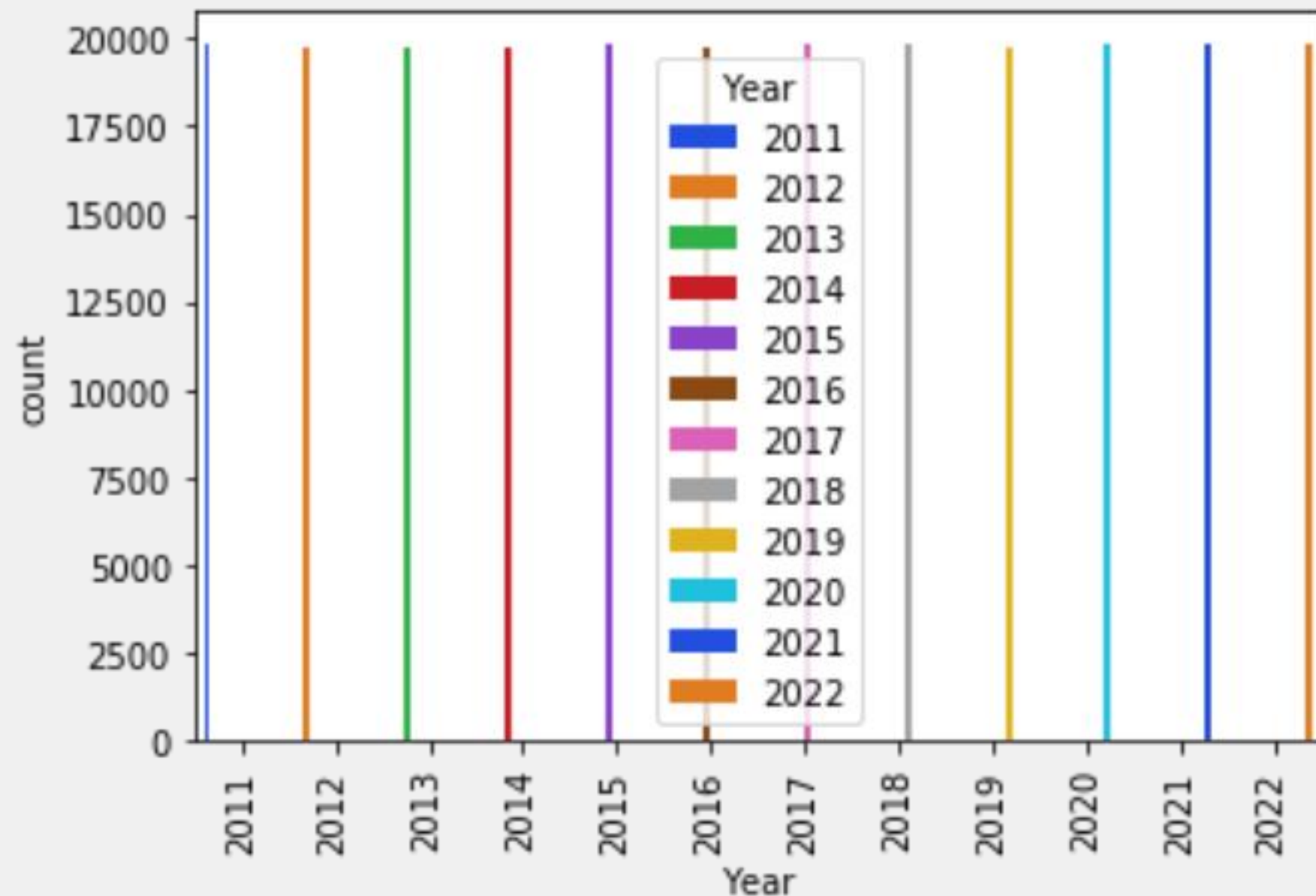
Q5.

Maximum deaths in different Years

#code

```
p = sns.countplot(x='Year', data = df, hue='Year', palette='bright')  
_ = plt.setp(p.get_xticklabels(), rotation=90)
```

The year 2021 has the highest death rates

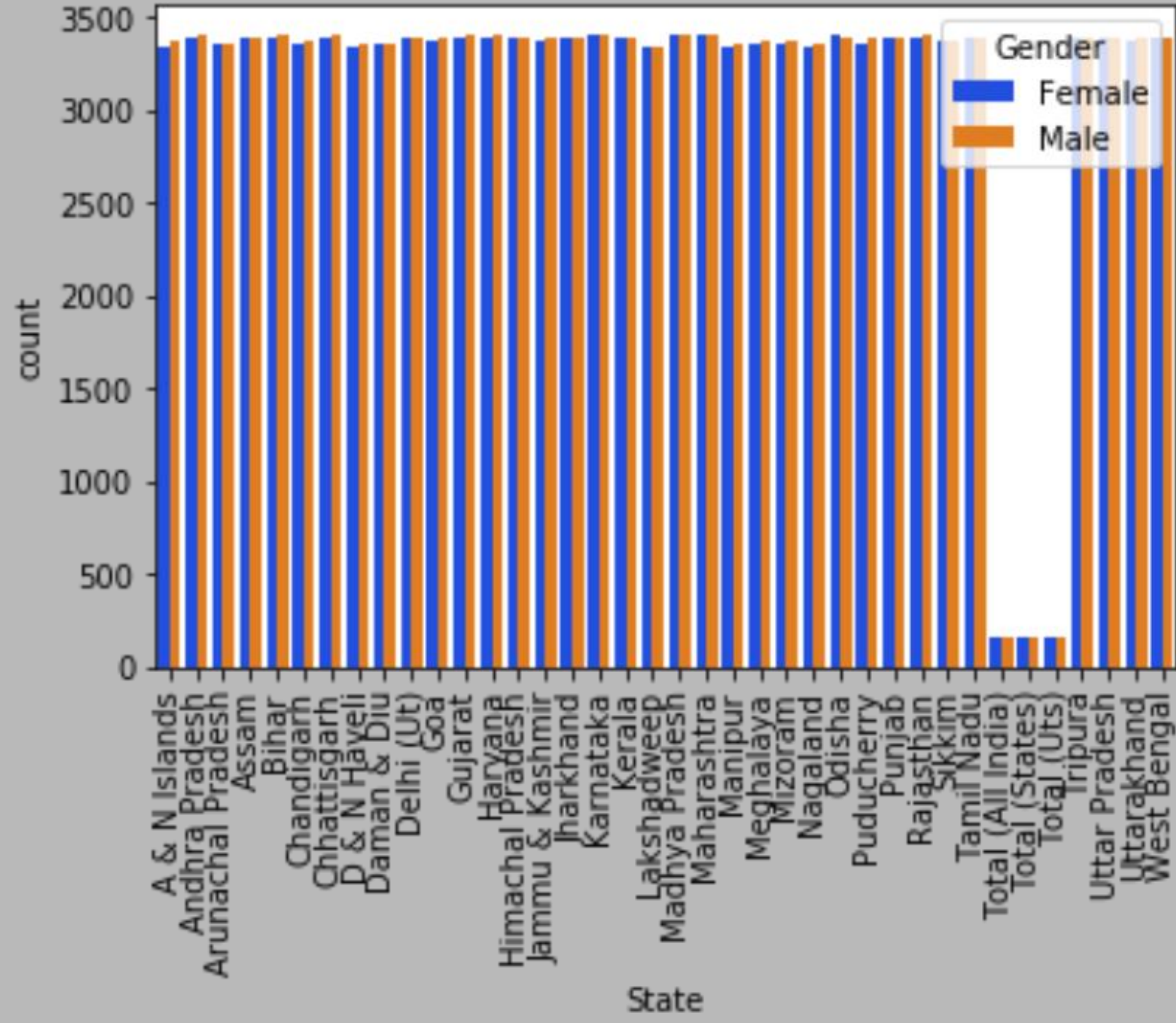


2021	19806
2015	19803
2022	19799
2018	19797
2011	19797
2017	19794
2020	19792
2012	19790
2019	19786
2016	19786
2013	19786
2014	19783

Name: Year, dtype: int64

Q6.
**Which State has highest death rates among Male
and Female**

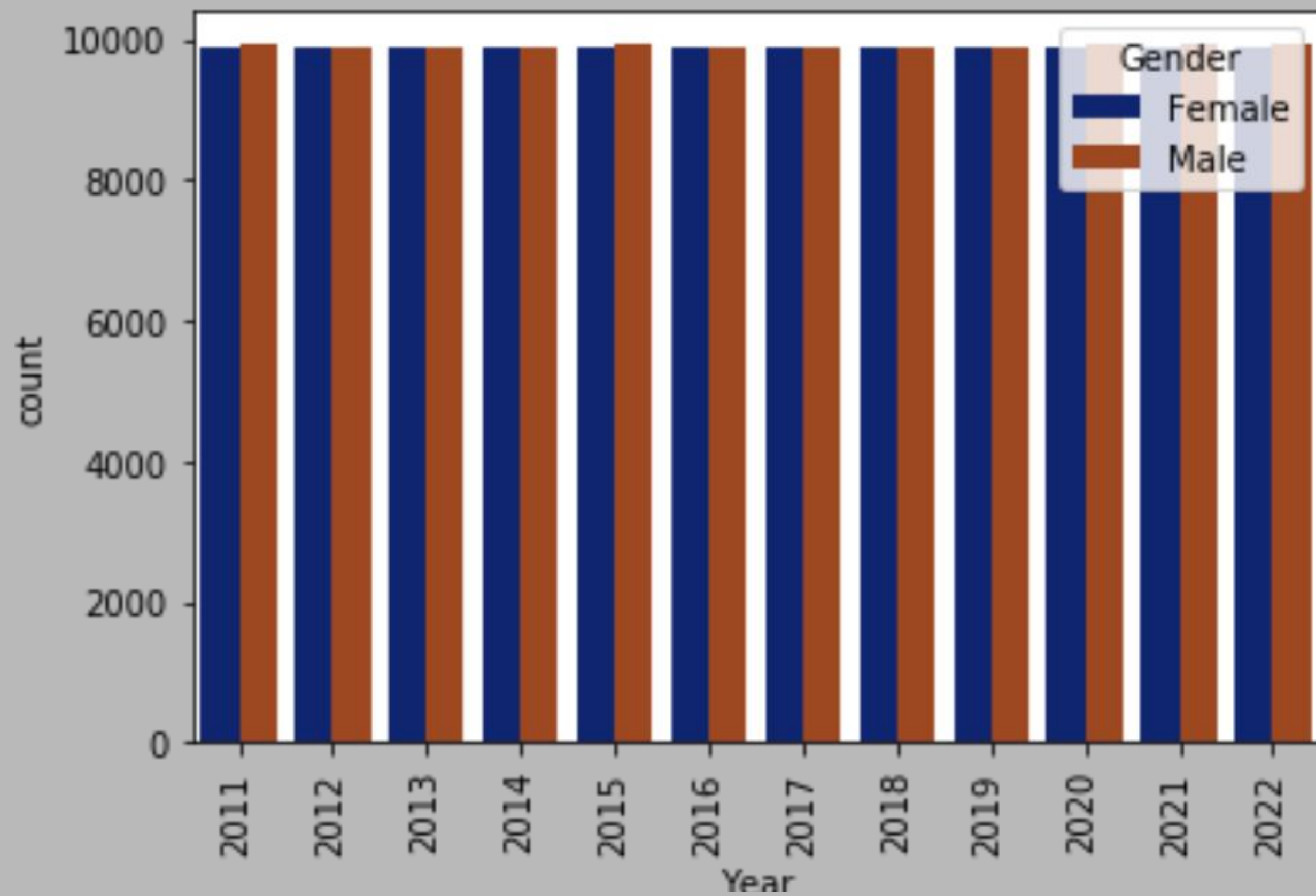






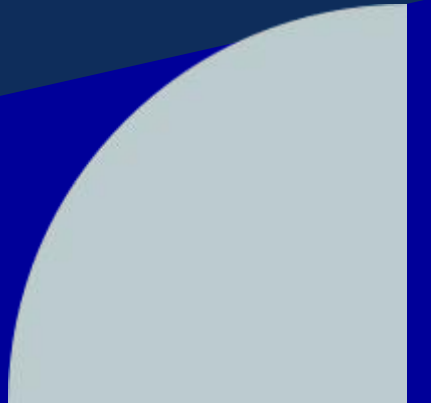
Q7.

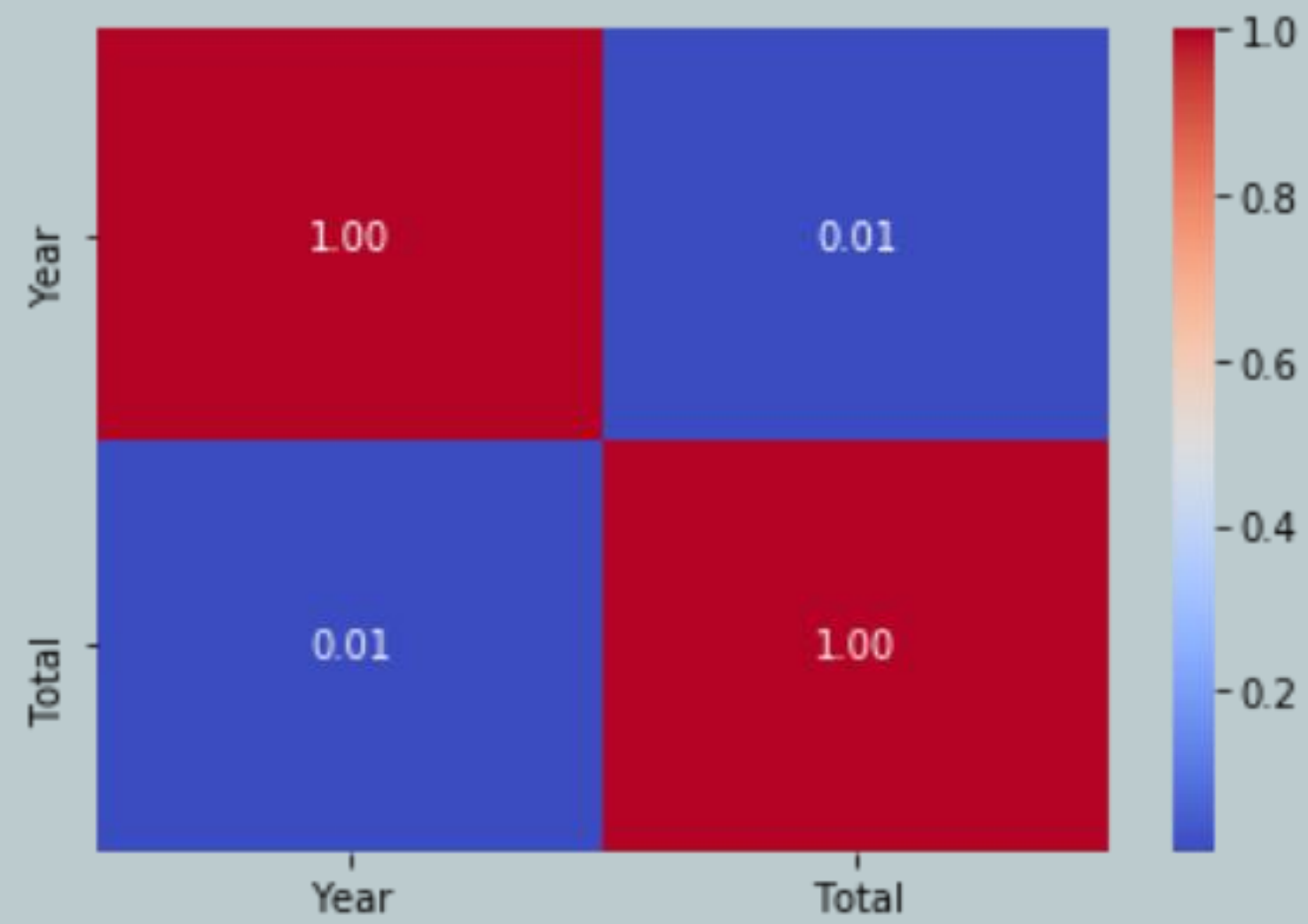
Which Year has the highest death rate among Male and Female.



Q8.

The correlation of
the dataset





CONCLUSION

- As observed above we can say that the maximum death is caused by 'OTHERS'
- The minimum death is caused by three means of disease
- As observed above we can say that the maximum death is in 3 States which is
 - >Maharashtra (6792)
 - >Madhya Pradesh (6792)
 - >Karnataka (6792)
- The maximum deaths among Male and Female is Male with rate of 118879
- Maximum Death rates is among the age group of 15-29
- The year 2021 has the highest death rates

THANK
YOU

