

Sample-2

Introduction: Tell me what problem you are working on? Why is this problem interesting and important. State specific research questions your group will work on. Introduce recent research done in area related to your problem. You can pack all this together to motivate us. Do keep it short, to the point, and interesting.

My research questions are:

Are there differences in mean anxiety levels of males and females between those in Europe in those in Asia?

Are there differences in mean anxiety levels of males and females between high SES countries and low SES countries?

These questions are valid, as research has shown that the burden of mental disease has increased globally in the past three decades and has varied across demographics such as gender (Piao et al., 2022). Therefore, it is a very relevant and important topic, and more investigation can be done as to how specific mental illnesses such as anxiety differ globally and between genders. Additionally, research has focused primarily on Western and high SES countries, and has left out other cultures and low SES countries (Baxter et al., 2013). It would be interesting for us to compare the results of Western culture (Europe) with more Eastern countries (Asia), and see if there is a difference, or if the results of existing studies are true for other areas of the world as well. It would also be interesting to see how anxiety levels differ between high SES countries, which have been studied, and low SES countries which have not been studied as much.

Data: Tell me about the data resource and explain dimensions of the data, variables in the data, and how does this data relate to your research questions.

My final dataset is a subset of the data from this source:

Saloni Dattani, Hannah Ritchie and Max Roser (2021) - "Mental Health". Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/mental-health' [Online Resource]

accessed on Kaggle:

<https://www.kaggle.com/datasets/thedevastator/gender-mental-disorder-prevalence-2019?select=prevalence-of-anxiety-disorders-males-vs-females.csv>

I used a subset of this data, focusing only on the anxiety data and not using the data for the other disorders included. The data shows the entity (country), country code, year, prevalence of anxiety for females (%), prevalence of anxiety for males (%), population, and continent. I cleaned up the dataset, which now looks like this in R:

	index	Entity	Code	Year	Population	Continent	Gender	Anxiety
1	1	Afghanistan	AFG	1990	12412311	Asia	MaleAnxiety	3.556843
2	1	Afghanistan	AFG	1990	12412311	Asia	FemaleAnxiety	5.971172
3	2	Afghanistan	AFG	1991	13299016	Asia	MaleAnxiety	3.548885
4	2	Afghanistan	AFG	1991	13299016	Asia	FemaleAnxiety	5.980482
5	3	Afghanistan	AFG	1992	14485543	Asia	MaleAnxiety	3.542779
6	3	Afghanistan	AFG	1992	14485543	Asia	FemaleAnxiety	5.988175
7	4	Afghanistan	AFG	1993	15816601	Asia	MaleAnxiety	3.538304
8	4	Afghanistan	AFG	1993	15816601	Asia	FemaleAnxiety	5.993858
9	5	Afghanistan	AFG	1994	17075728	Asia	MaleAnxiety	3.535309
10	5	Afghanistan	AFG	1994	17075728	Asia	FemaleAnxiety	5.997363
11	6	Afghanistan	AFG	1995	18110662	Asia	MaleAnxiety	3.533797
12	6	Afghanistan	AFG	1995	18110662	Asia	FemaleAnxiety	5.998540
13	7	Afghanistan	AFG	1996	18853444	Asia	MaleAnxiety	3.535415
14	7	Afghanistan	AFG	1996	18853444	Asia	FemaleAnxiety	5.996443

Showing 1 to 15 of 5,040 entries, 8 total columns

Exploratory Data Analysis:

Numerical summary for Male and Female Anxiety based on continent:

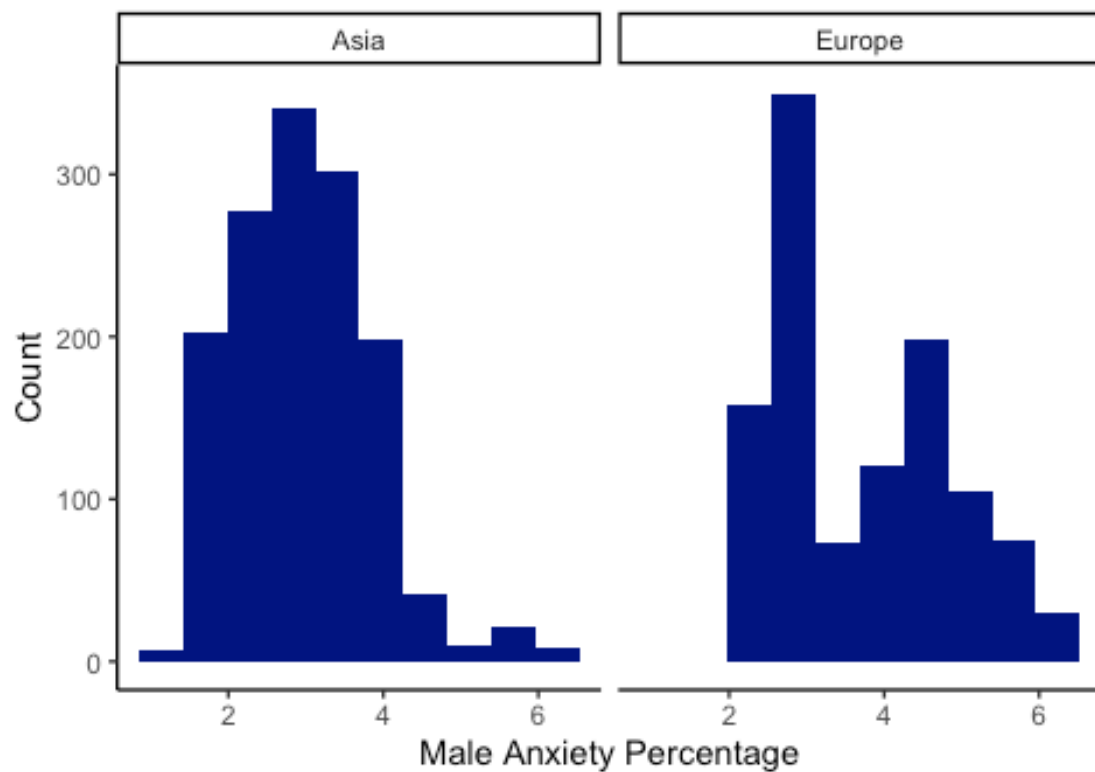
```
AnxietyEurAis2 %>% group_by(Continent) %>% summarise(MeanMaleAnxiety=mean(Male),
MeanFemaleAnxiety=mean(Female),
sdMaleAnxiety=sd(Male),
sdFemaleAnxiety=sd(Female))

## # A tibble: 2 × 5
##   Continent MeanMaleAnxiety MeanFemaleAnxiety sdMaleAnxiety sdFemaleAnxiety
##   <chr>          <dbl>          <dbl>          <dbl>          <dbl>
## 1 Asia           2.98           4.90           0.894           1.55
## 2 Europe         3.71           6.39           1.15           2.01
```

Histograms:

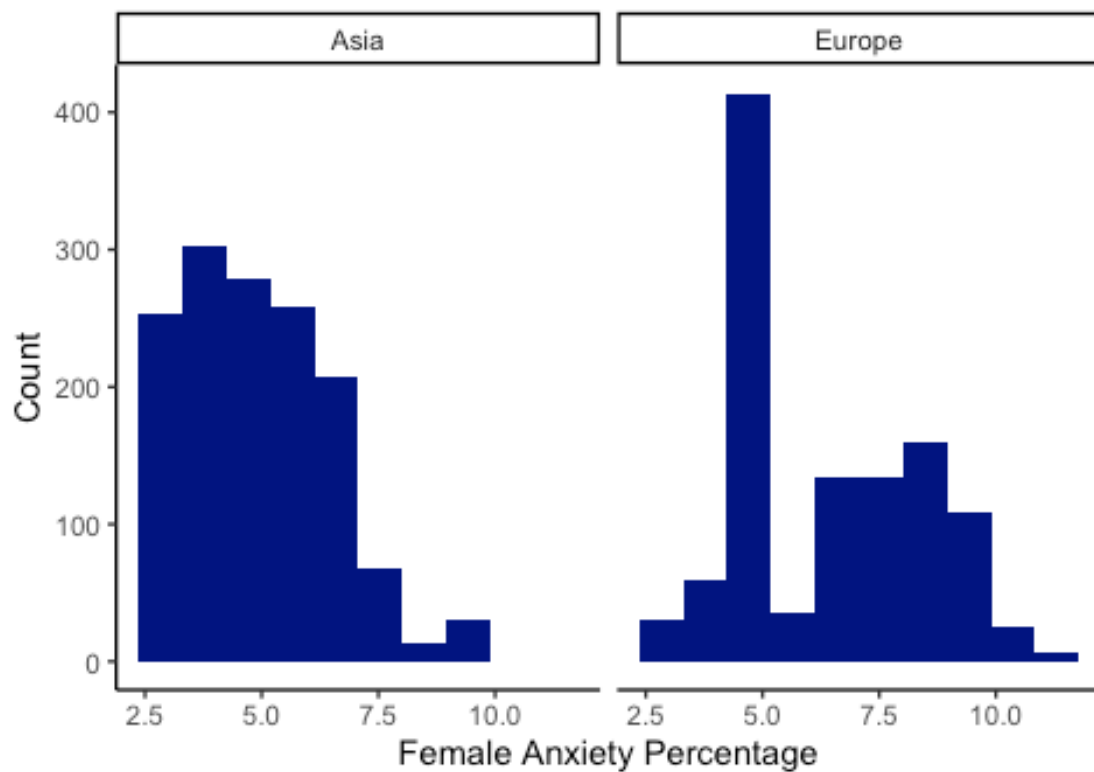
```
AnxietyEurAis2 %>% ggplot(aes(x=Male))+geom_histogram(bins=10,fill="navy")+facet_wrap(~Continent)+
  theme_classic()+xlab("Male Anxiety Percentage")+ylab("Count")+
  ggtitle("Fig 1: Distributions of Male Anxiety Percentage", "Asia vs Europe")
```

Fig 1: Distributions of Male Anxiety Percentage
Asia vs Europe



```
AnxietyEurAis2 %>% ggplot(aes(x=Female))+geom_histogram(bins=10,fill="navy")+
facet_wrap(~Continent)+
  theme_classic()+xlab("Female Anxiety Percentage")+ylab("Count")+
  ggtitle("Fig 2: Distributions of Female Anxiety Percentage", "Asia vs Europe")
```

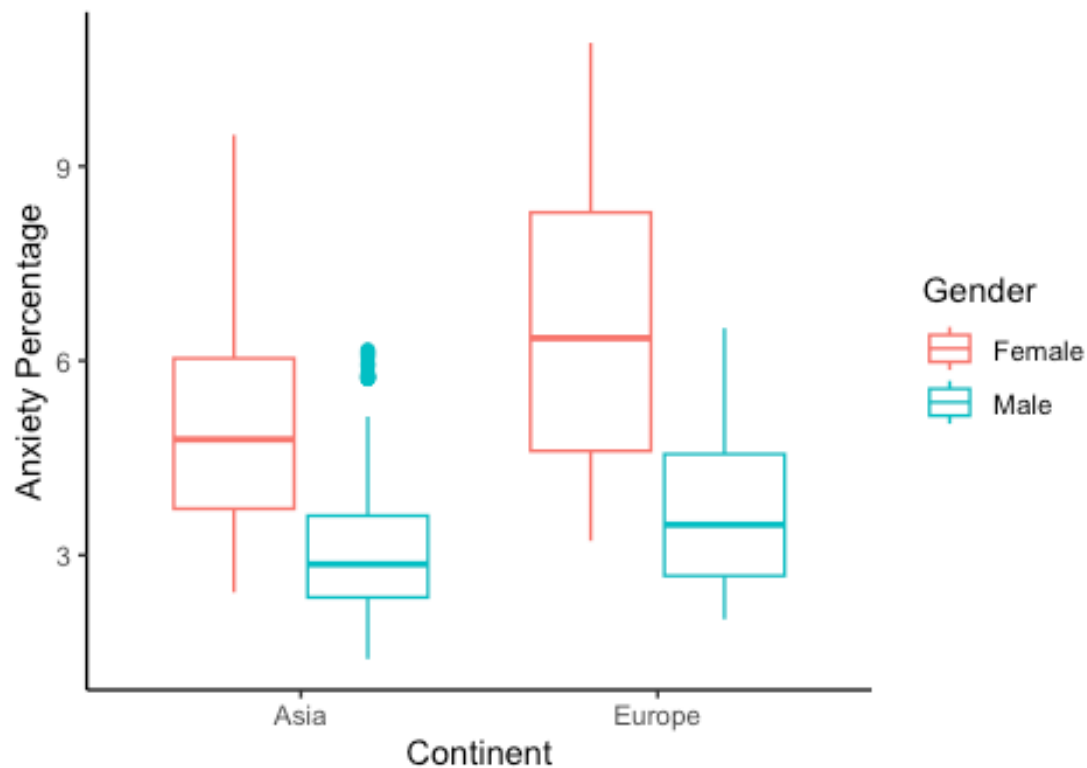
Fig 2: Distributions of Female Anxiety Percentage
Asia vs Europe



Boxplot:

```
AnxietyEurAis3 %>%  
  ggplot(aes(x=Continent,y=Anxiety,col=Gender))+geom_boxplot()+theme_classic(  
) +  
  ggtitle("Fig 3. Anxiety Percentage by Continent (Asia vs Europe)", "Colored  
by Gender") +  
  ylab("Anxiety Percentage")
```

Fig 3. Anxiety Percentage by Continent (Asia vs Europe)
Colored by Gender



The histograms show the distributions of male and female anxiety for Europe and Asia. Both distributions for Asia appear right sd. Both distributions for Europe appear bimodal and right skewed. The boxplot and summary statistics show that female anxiety percentage has a higher central tendency and a higher standard deviation than male anxiety percentage. They also show that anxiety percentage in Europe has a slightly higher central tendency and standard deviation than that of Asia.

From what we have seen in the EDA, it seems that there are differences in average anxiety levels between different continents and different genders. From the histograms and boxplots, you can see that female anxiety percentage tends to be much higher than that of males. This may have to do with male anxiety being underreported, but there may really be gender differences in anxiety. Additionally, looking at our boxplots and summary statistics, it appears that anxiety percentage in Europe tends to be slightly higher on average than that of Asia. This may have to do with cultural differences and is interesting because Europe has been heavily studied in terms of mental illness, but this suggests that Europe may not be representative of anxiety levels in other areas of the globe. I need to do more analysis to determine if there is a real difference.