Dataset consisting of: Body girth measurements and skeletal diameter measurements, as well as age, weight, height and gender

importing packages

In [118…
```
import pandas as pd
import seaborn as sns
from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt
```

Read the data

In [119…
```
data = pd.read_csv("bdims.csv")
```

Separating the data for individuals over 21 years of age and male vs female

In [120…
```
data_over_21 = data[data["age"] >= 21]
maled = data_over_21[data_over_21["sex"] == 1]
femaled = data_over_21[data_over_21["sex"] == 0]
```

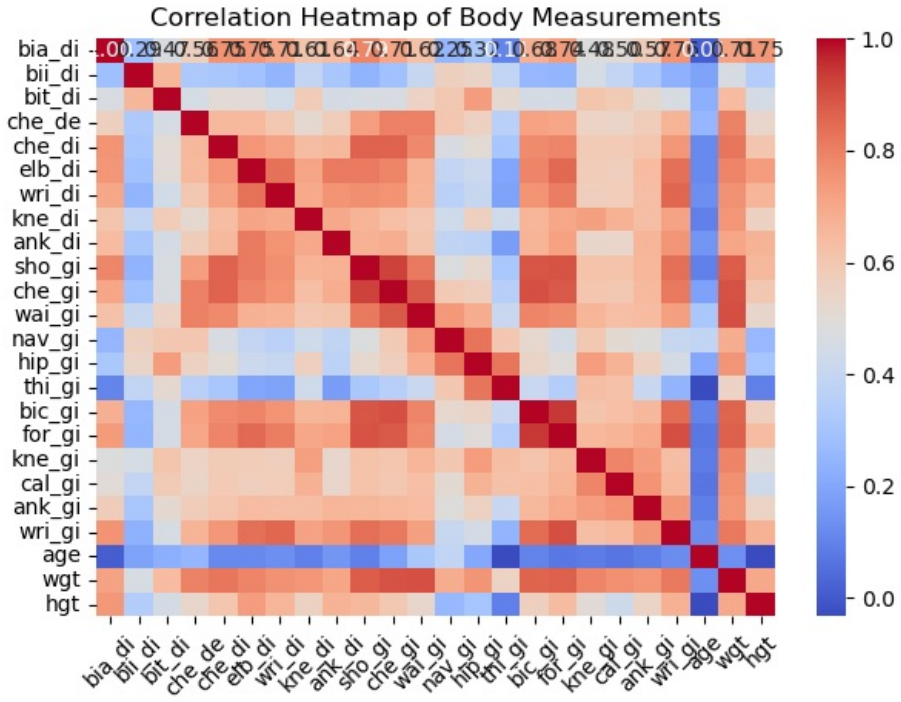
Looking at an overview of average height and weight

In [121…
```
print(f'Average male height: {maled["hgt"].mean():.1f} cm\nAverage male weight: {maled["wgt"].mean():.1f} kg')
print(f'Average female height: {femaled["hgt"].mean():.1f} cm\nAverage female weight: {femaled["wgt"].mean():.1
```

```
Average male height: 178.0 cm
Average male weight: 78.5 kg
Average female height: 165.2 cm
Average female weight: 61.1 kg
```

Generate heatmap correlation matrix for numeric values

In [122…
```
data_over_21wos = data_over_21.drop(columns=['sex'])
correlation_matrix = data_over_21wos.corr()
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt=".2f")
plt.title('Correlation Heatmap of Body Measurements')
plt.xticks(rotation=45)
plt.yticks(rotation=0)
plt.tight_layout()
plt.show()
```

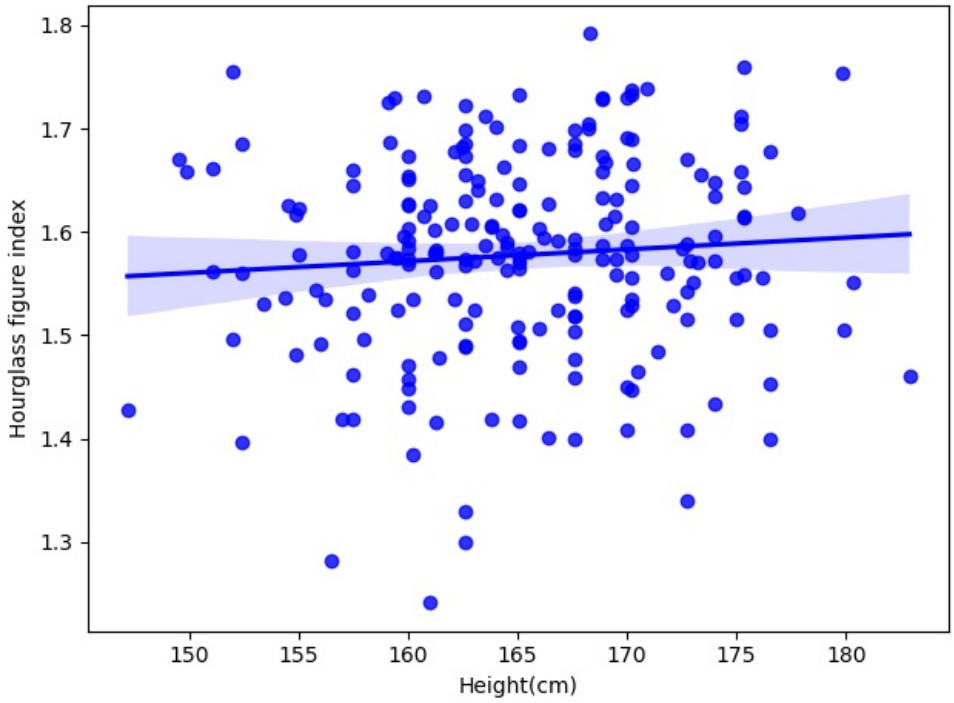


Now i will check whether if there is a correlation between a females height, weight and age with their body shape replicating an hourglass figure

In [123…
```
female_weight = femaled["wgt"]
hg = (femaled["hip_gi"]*1.25+femaled["sho_gi"])/(2*femaled["wai_gi"]) #formula used for hourglass shape is (hip
```

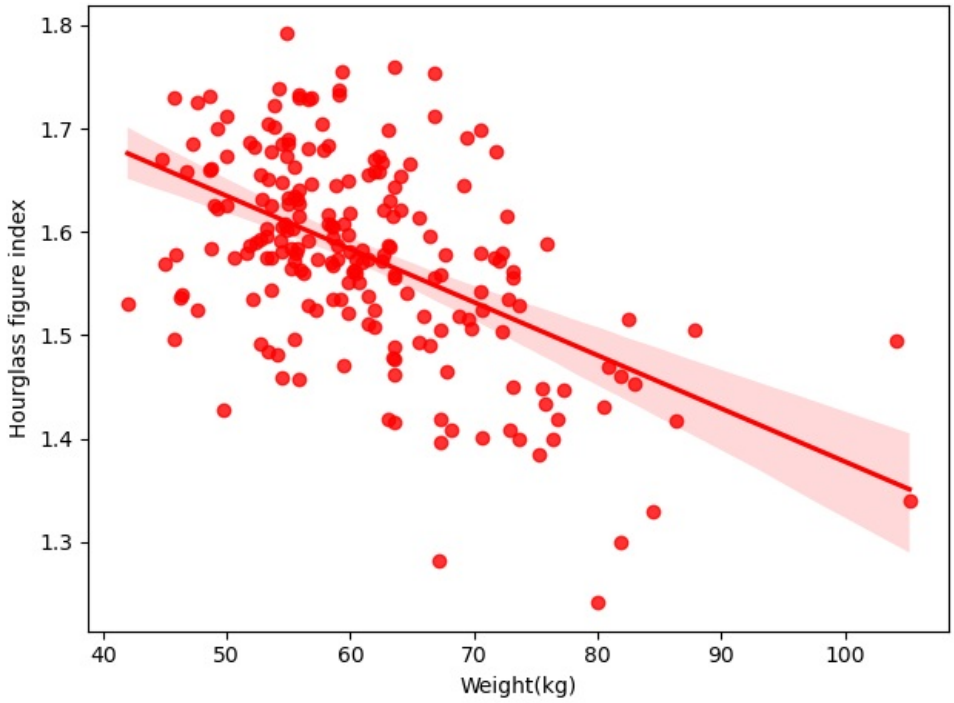
In [124…
```
sns.regplot(x=female_height, y=hg, color="blue")
```

```
plt.xlabel(f"Height(cm)")
plt.ylabel("Hourglass figure index")
plt.tight_layout()
plt.show()
```



There seems to be no correlation

In [125…
```
fw = femaled["wgt"]
sns.regplot(x=fw, y=hg, color="red")
plt.xlabel(f"Weight(kg)")
plt.ylabel("Hourglass figure index")
plt.tight_layout()
plt.show()
```



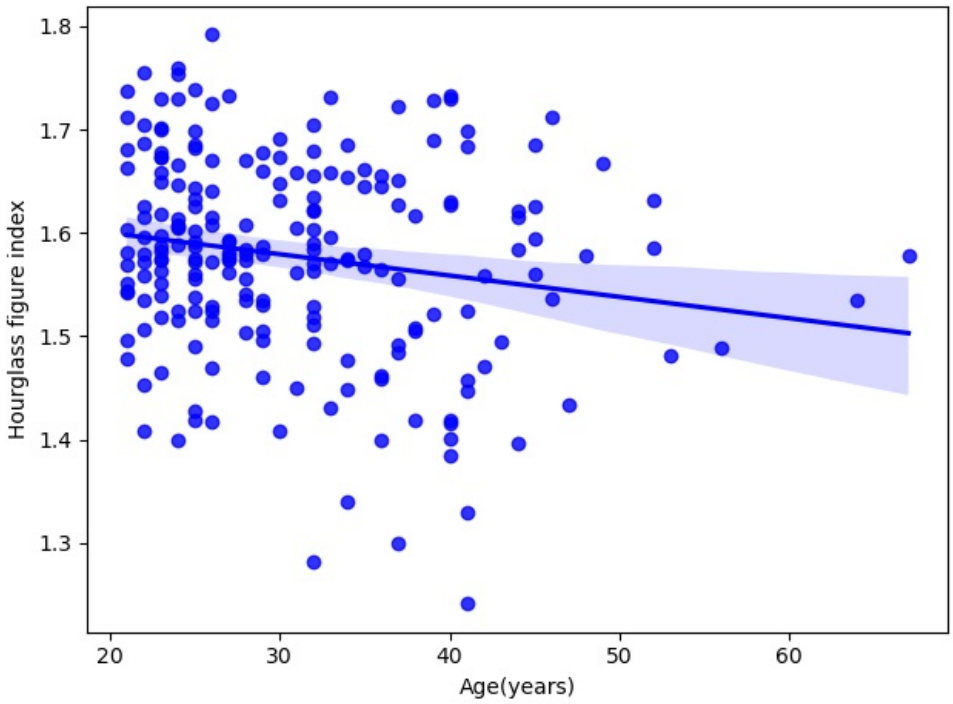
There seems to be a clear negative correlation

In [126…
```
corr1 = fw.corr(hg)
print(f"Correlation coefficient: {corr1:.2f}")
```

```
Correlation coefficient: -0.52
```

In [127…
```
fage = femaled["age"]
sns.regplot(x=fage, y=hg, color="blue")
plt.xlabel(f"Age(years)")
plt.ylabel("Hourglass figure index")
plt.tight_layout()
```

```
plt.show()
```



Very slight negative correlation

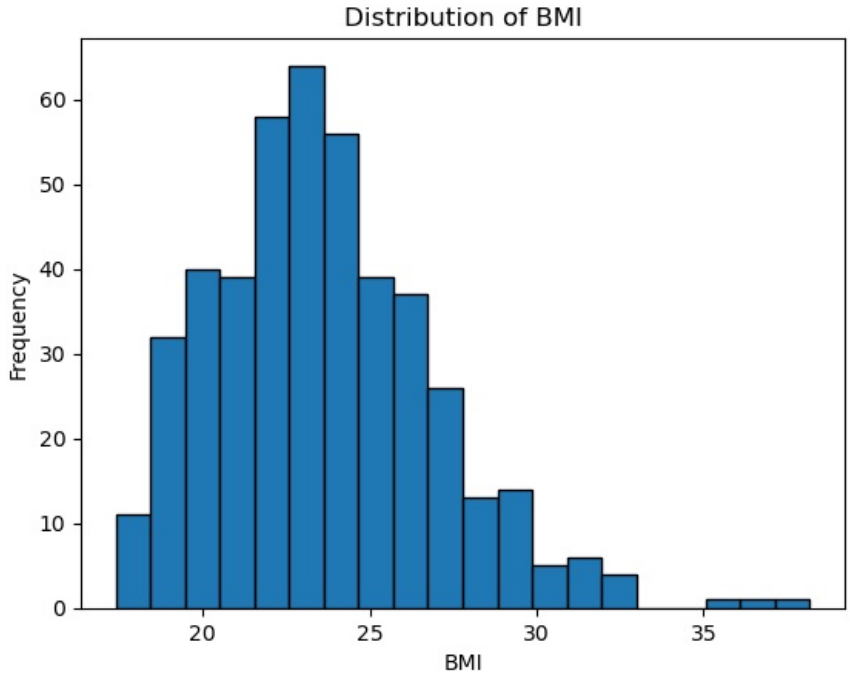
In [128…
```
corr1 = fage.corr(hg)
print(f"Correlation coefficient: {corr1:.2f}")
```

```
Correlation coefficient: -0.18
```

Analyzing BMI values

Distribution of BMI

In [129…
```
plt.hist(gBMI, bins=20, edgecolor='black')
plt.xlabel('BMI')
plt.ylabel('Frequency')
plt.title('Distribution of BMI')
plt.show()
```



BMI and age

In [130…
```
gBMI = data_over_21["wgt"]/((data_over_21["hgt"]/100)**2)
gAge = data_over_21["age"]
print(f"BMI and age correlation: {gBMI.corr(gAge):.2f}")
```

```
BMI and age correlation: 0.21
```

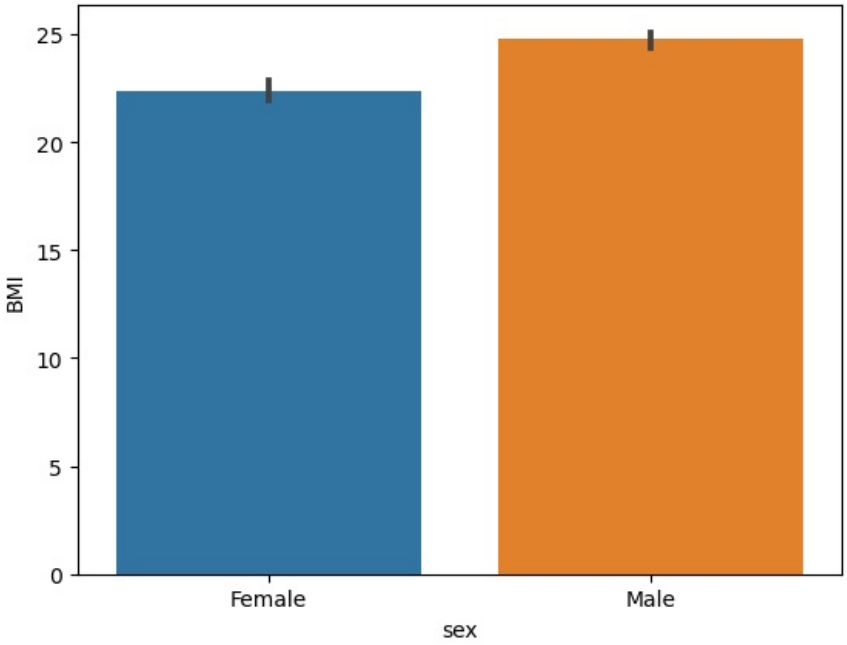
Male BMI and height

In [131…
```
mBMI = maled["wgt"]/((maled["hgt"]/100)**2)
mAge = maled["age"]
print(f"BMI and age correlation: {mBMI.corr(mAge):.2f}")
```

```
BMI and age correlation: 0.20
```

BMI and sex

In [132…
```
sns.barplot(x=data_over_21["sex"], y=gBMI)
plt.xticks(ticks=[0, 1], labels=['Female', 'Male'])
plt.ylabel("BMI")
plt.show()
```



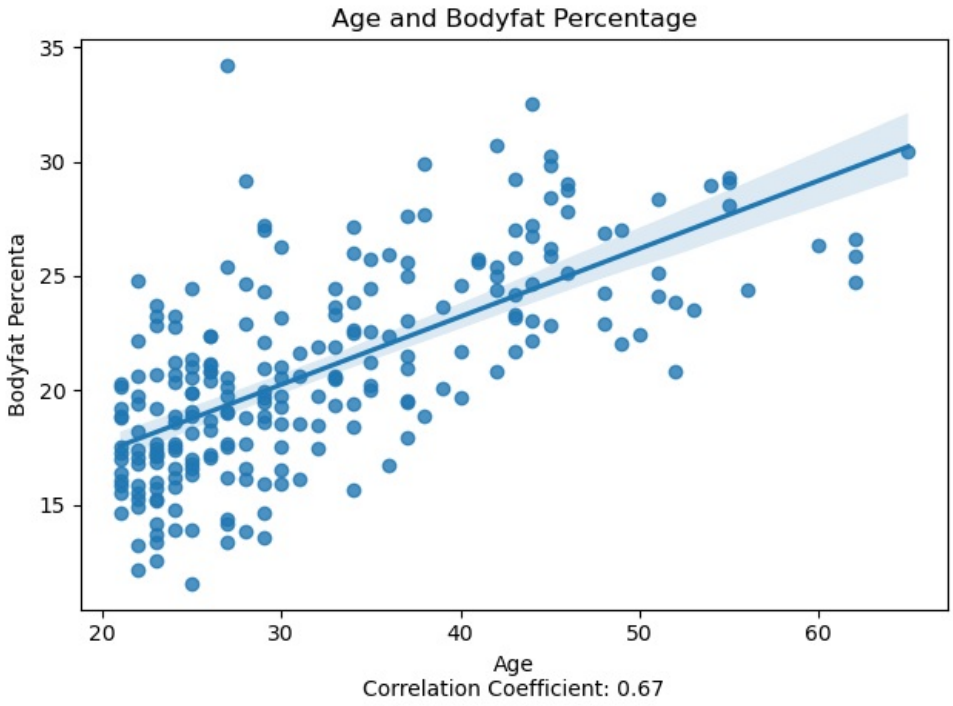
Bodyfat analysis

Male bodyfat percentage estimate

In [133…
```
mbodyfat = (1.20 * mBMI) + (0.23 * mAge) - 16.2
print("Average bodyfat percentage: ", end="")
print(mbodyfat.mean())
```

```
Average bodyfat percentage: 20.928107434850016
```

In [134…
```
sns.regplot(x=mAge, y=mbodyfat)
a = mAge.corr(mbodyfat)
plt.ylabel("Bodyfat Percenta")
plt.xlabel(f"Age\nCorrelation Coefficient: {a:.2f}")
plt.title("Age and Bodyfat Percentage")
plt.tight_layout()
plt.show()
```



Wrist and biceps girth

In [135…

```
biceps = maled["bic_gi"]
wrist = maled["wri_gi"]
sns.regplot(x=wrist, y=biceps)
b = wrist.corr(biceps)
plt.xlabel(f"Wrist circumference(cm)\nCorrelation coefficient: {b:.2f}")
plt.ylabel("Biceps circumference(cm)")
plt.tight_layout()
plt.show()
```