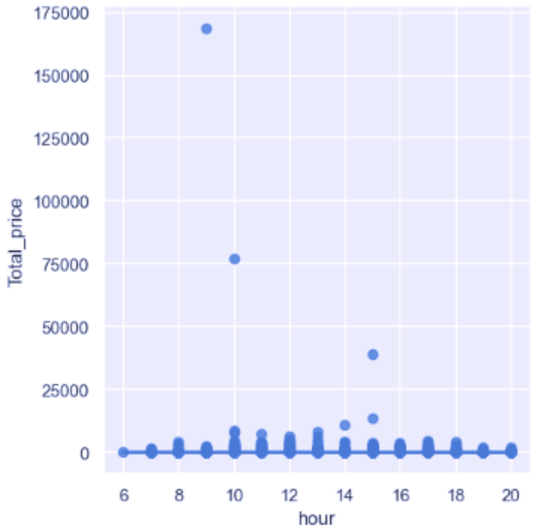
hypothesis Testing

Hypothesis

My insights says the the best selling time in the whole world is from 9 am to 4 o clock

```
sns.lmplot(x='hour', y ="Total_price", data=df);
```



let's create two sets of data

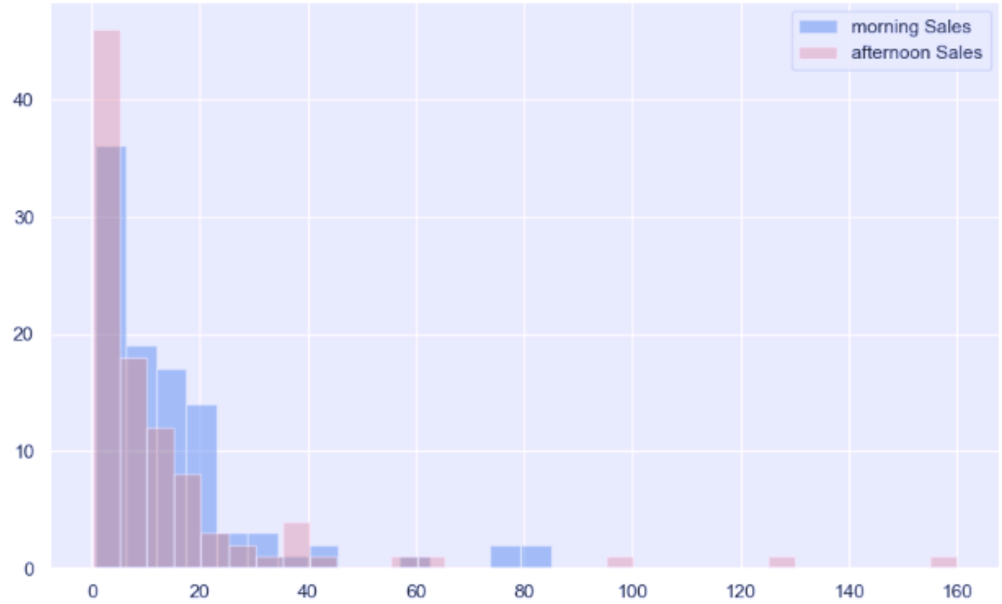
morning Sales : If time is from 8 to 17

afternoon Sales: If time is from 18 to 7

Take a sample of 100 record in each dat

```
Set alpha (set alpha = 0.05)
```

```
'H0': "Total price does not increase for Sale",
'H1': "Total price increases for Large Sale"
```



Variance

```
# ANSWER
print("High sales:", part1.var())
print("Low sales:", part2.var())
```

```
High sales: 285.81379256565657
Low sales: 566.5941854141413
```

Calculate Standard Deviation

```
std = np.sqrt(( part1.var() +  part2.var()) / 2)
print('standard deviation is:', std)
```

```
standard deviation is: 20.644708498545068
```

Calculate test statistic

```
t = (part1.mean() - part2.mean()) /(std *np.sqrt(2/N))
print('t:', t)
```

```
t: 0.23311391163962647
```

Calculate Degree of Freedom

```
df = 2 * N - 2
df
```

```
198
```

Find the p-value

```
p = 1 - stats.t.cdf(abs(t), df=df)
print("t = " + str(t))
print("p = " + str(2 * p))
```

```
t = 0.23311391163962647
p = 0.815913764073509
```

Summary

- There is no difference between morning sales and afternoon sales when we look at the global market of this business
- Considering analysing each country would possibly highlight different results.

```
t = 0.2331139116396264 7
p = 0.815913764073509
We fail to reject our null hypothesis.
Total price does not increase for Sale
```

Hourly sale time in the data