# Session8-Pandas(Handling with Outliers)

DAwithPython S8
Training Clarusway
Pear Deck - May 7, 2022 at 10:11AM

## Part 1 - Summary

Use this space to summarize your thoughts on the lesson

## Part 2 - Responses

Slide 1


Use this space to take notes:

## Slide 2


Use this space to take notes:

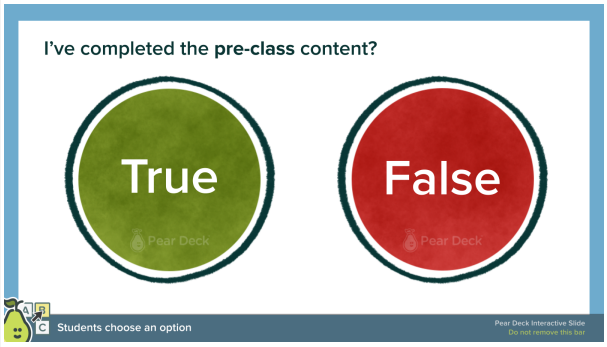
## Slide 3

### Table of Contents

- What is the Outliers?
- Detecting Outliers
- Handling with Outliers
- Some Useful Methods

Use this space to take notes:

| Slide 4 | Your Response |
| --- | --- |

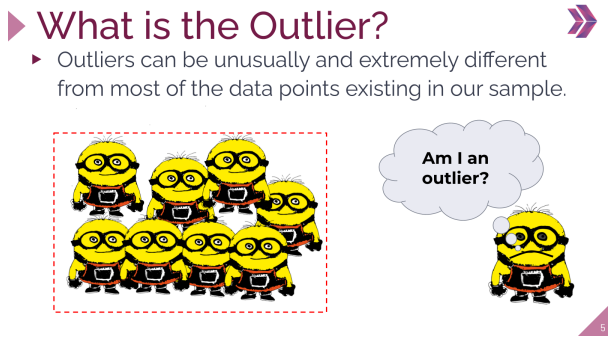



You Chose
- **False**

Other Choices
- True

Use this space to take notes:

## Slide 5




Use this space to take notes:

## Slide 6

## What is the Outlier?

- Outliers can create biased results while calculating the stats of the data due to its extreme nature, thereby affecting further statistical/ML models.

| Index | car_price |
|---|---|
| 1 | 22.000 |
| 2 | 24.000 |
| 3 | 1050 |
| 4 | 28.000 |
| 5 | 149.000 |

The abnormal values of given variable (`car_price`)

Such values are called ***outliers***

6

Use this space to take notes:

## Slide 7

## What is the Outlier?

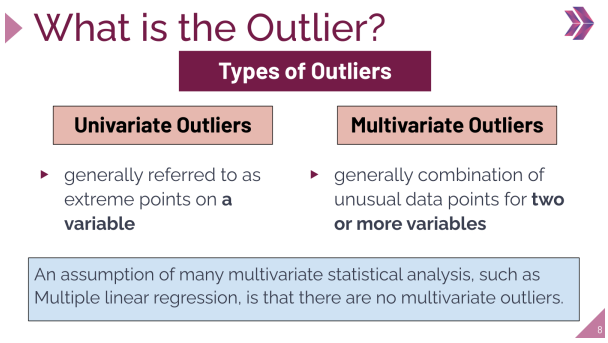
**Causes of Outliers**

- Data entries errors
- Measurement errors or instrument errors
- Sampling errors
- Data processing errors
- Natural novelties in data

7

Use this space to take notes:

## Slide 8

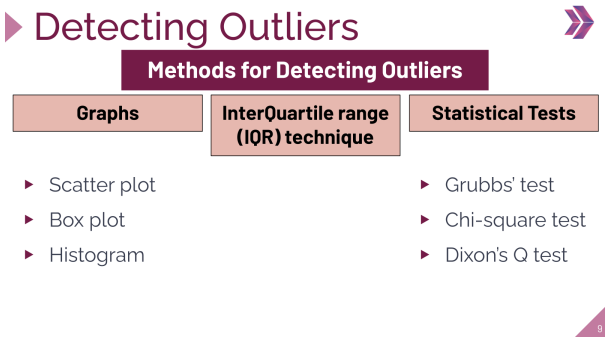



Link(s) on this slide:

- https://en.wikiversity.org/wiki/Multivariate_statistics

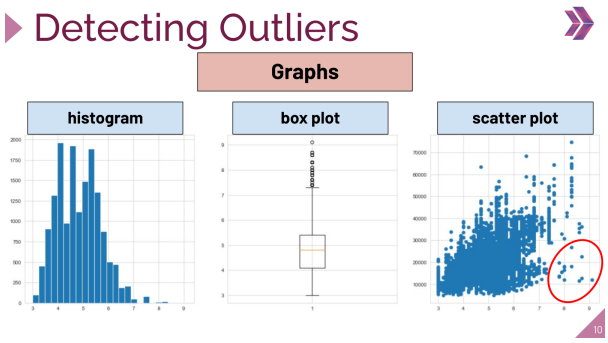
Use this space to take notes:

## Slide 9



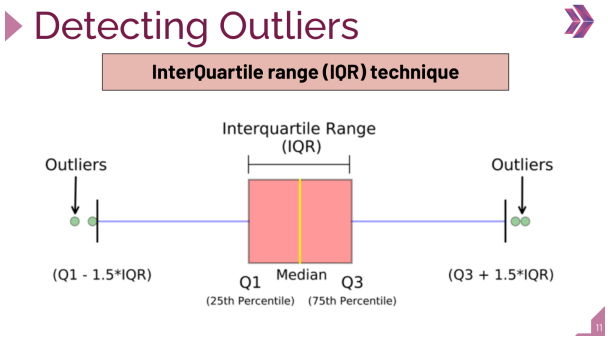

Use this space to take notes:

## Slide 10



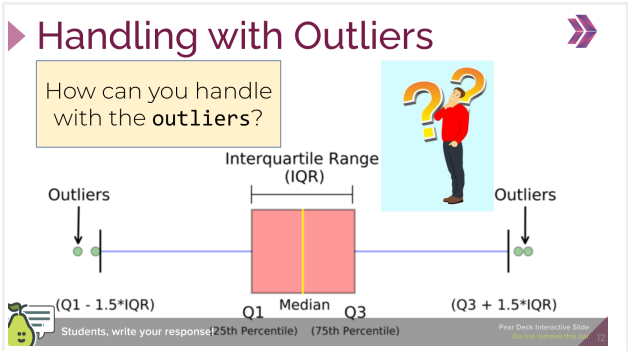

Use this space to take notes:

## Slide 11



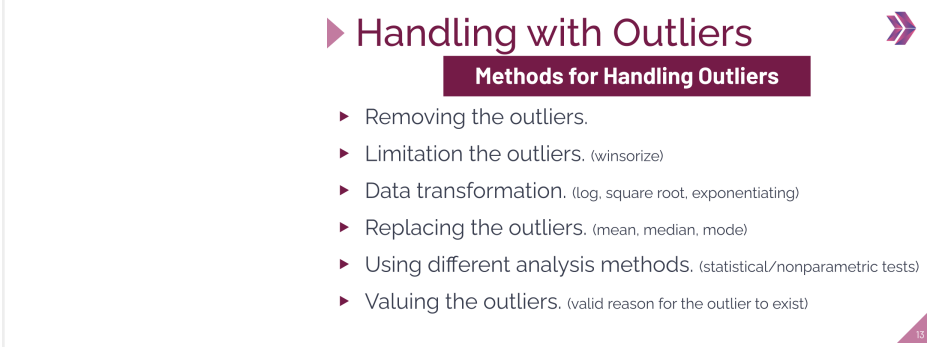

Use this space to take notes:

| Slide 12 | Your Response |
| --- | --- |
| | |




Use this space to take notes:

## Slide 13




Use this space to take notes:

## Slide 14

## Handling with Outliers

**Guideline for Handling Outliers**

If the outlier in question is:

- A measurement error or data entry error, correct the error if possible. If you can't fix it, remove that observation because you know it's incorrect.
- Not a part of the population you are studying (i.e., unusual properties or conditions), you can legitimately remove the outlier.
- A natural part of the population you are studying, you should not remove it.

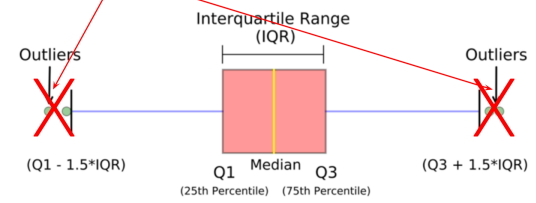
14

Use this space to take notes:

## Slide 15

## Handling with Outliers

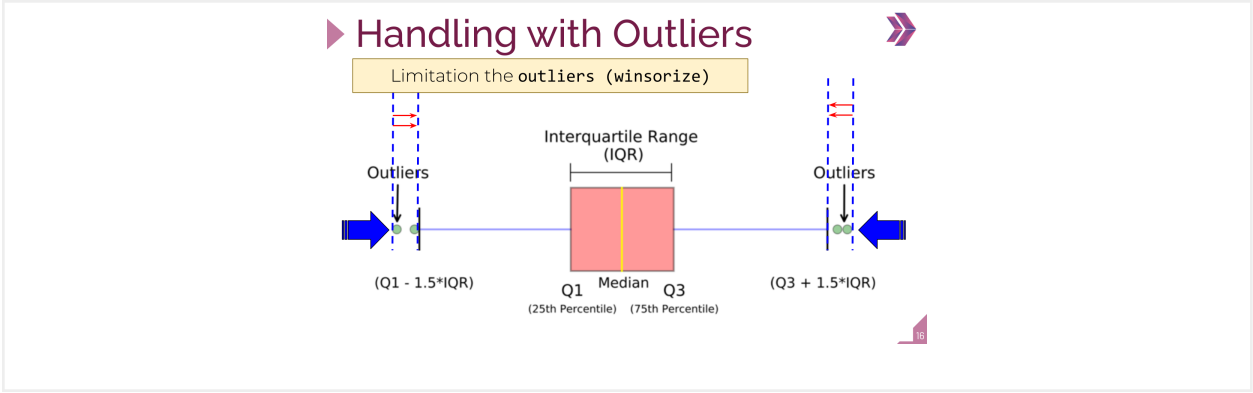
Removing the `outliers`




15

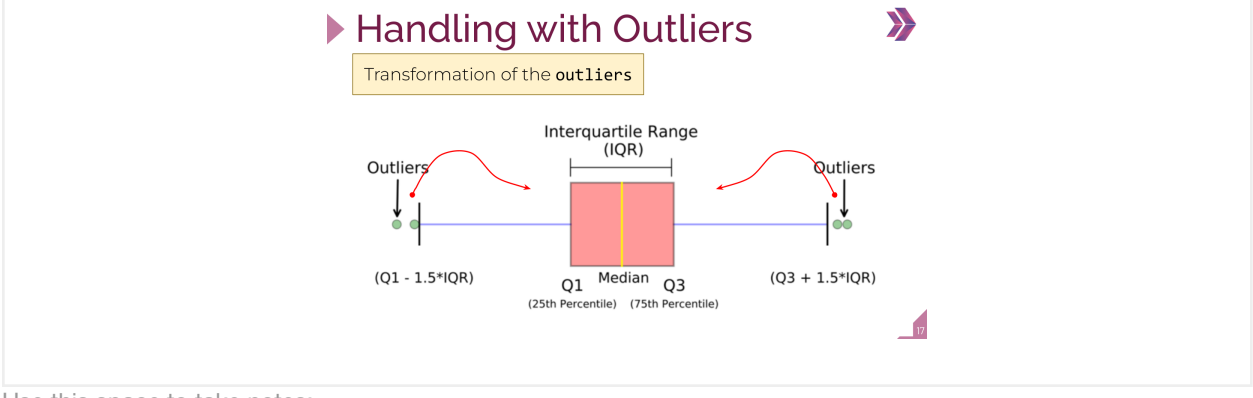
Use this space to take notes:

## Slide 16




Use this space to take notes:

## Slide 17

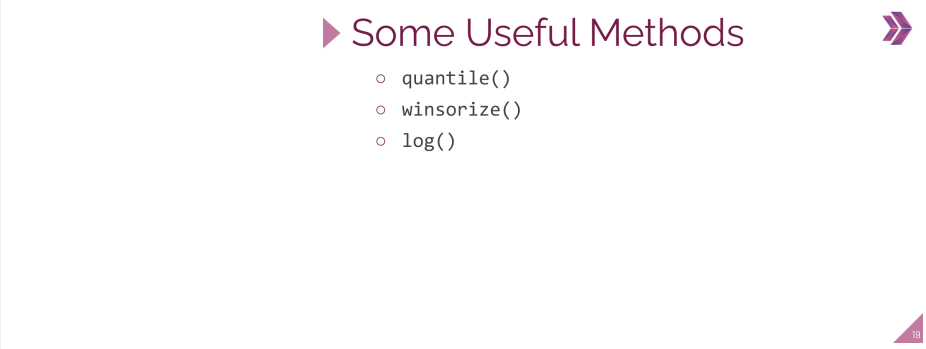



Use this space to take notes:

| Slide 18 | Your Response |
| --- | --- |
| | |


Use this space to take notes:

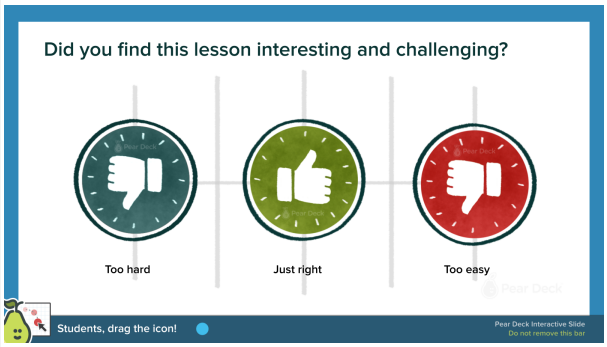
## Slide 19




Use this space to take notes:

## Slide 20


Use this space to take notes:

| Slide 21 | Your Response |
| --- | --- |
|  |  |



Use this space to take notes:

| Slide 22 |
| --- |

# THANKS!

## Any questions?

You can find us at:

steve_w@clarusway.com
michael_g@clarusway.com



Use this space to take notes: