Calculating Descriptive Statistics

Matthew Renze

@matthewrenze | www.matthewrenze.com

Overview

Introduction

Types of Analysis

Guidance

Demo

Descriptive Statistics

Describe data

Summarize data

Movie Runtime	
Statistic	Value (minutes)
Minimum	38
1st Quartile	93
Median	101
Mean	104
3rd Quartile	113
Maximum	219

Statistical Terms

Observations

Variables

Qualitative variable

Quantitative variable

ID	Date	Customer	Product	Quantity
1	2015-08-27	John	Pizza	2
2	2015-08-27	John	Soda	2
3	2015-08-27	Jill	Salad	1
4	2015-08-27	Jill	Milk	1
5	2015-08-28	Miko	Pizza	3
6	2015-08-28	Miko	Soda	2
7	2015-08-28	Sam	Pizza	1
8	2015-08-28	Sam	Milk	1

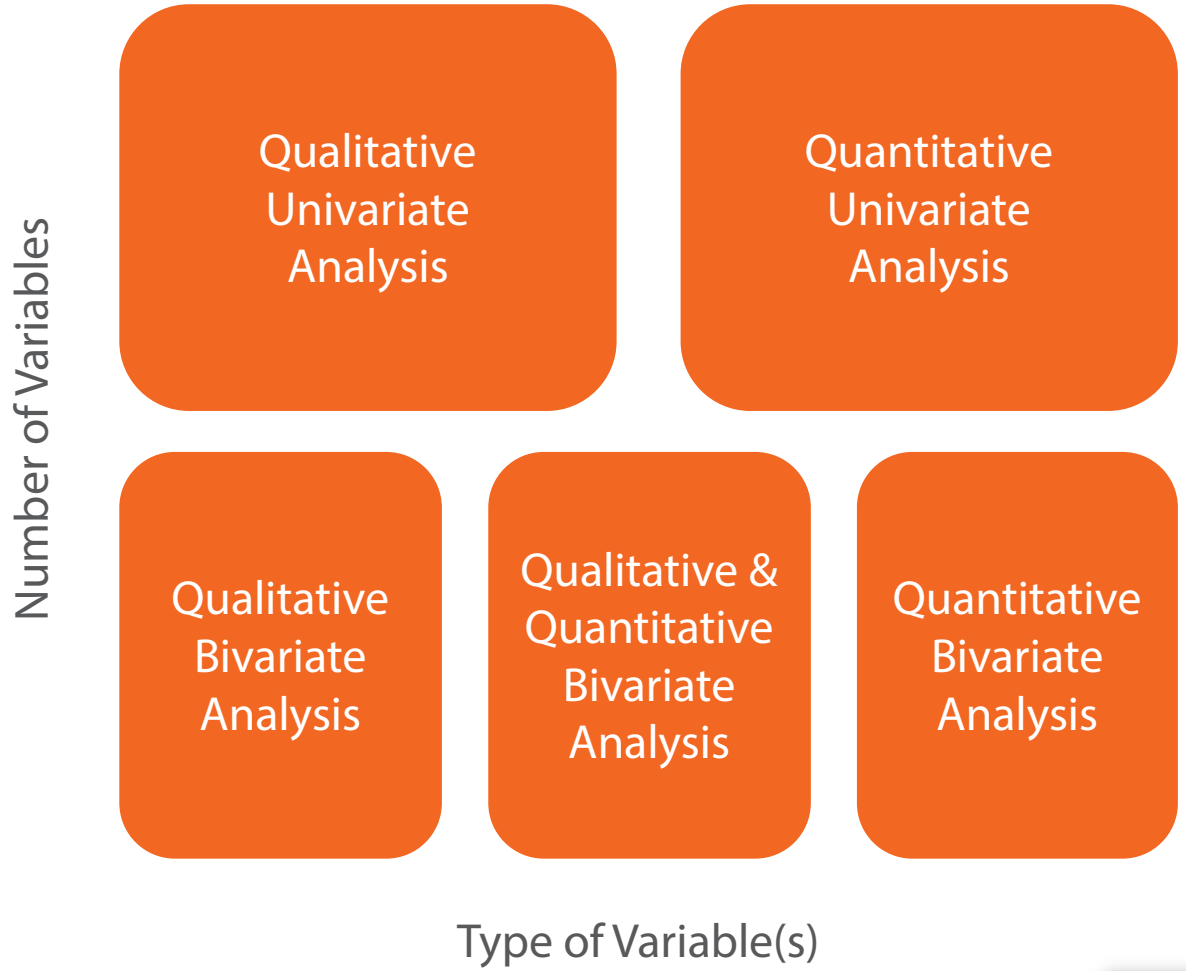
Types of Analysis

Number of variables

- Univariate
- Bivariate

Type of variables

- Qualitative
- Quantitative



Qualitative Univariate Analysis

Frequency

Percentage

Mode

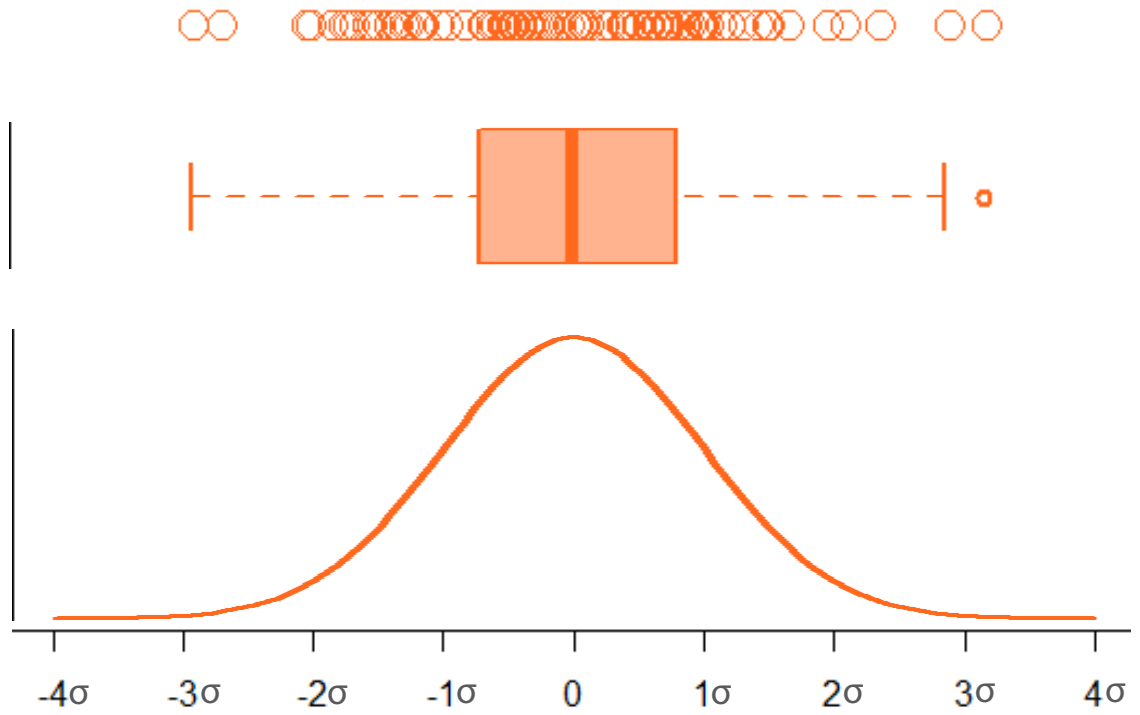
Movies by Genre		
Genre	Frequency	Percentage
Action	612	9%
Adventure	496	7%
Animation	168	2%
Comedy	1281	18%
Drama	1570	22%
Horror	269	4%
…	…	…

Quantitative Univariate Analysis

Location

Spread

Shape

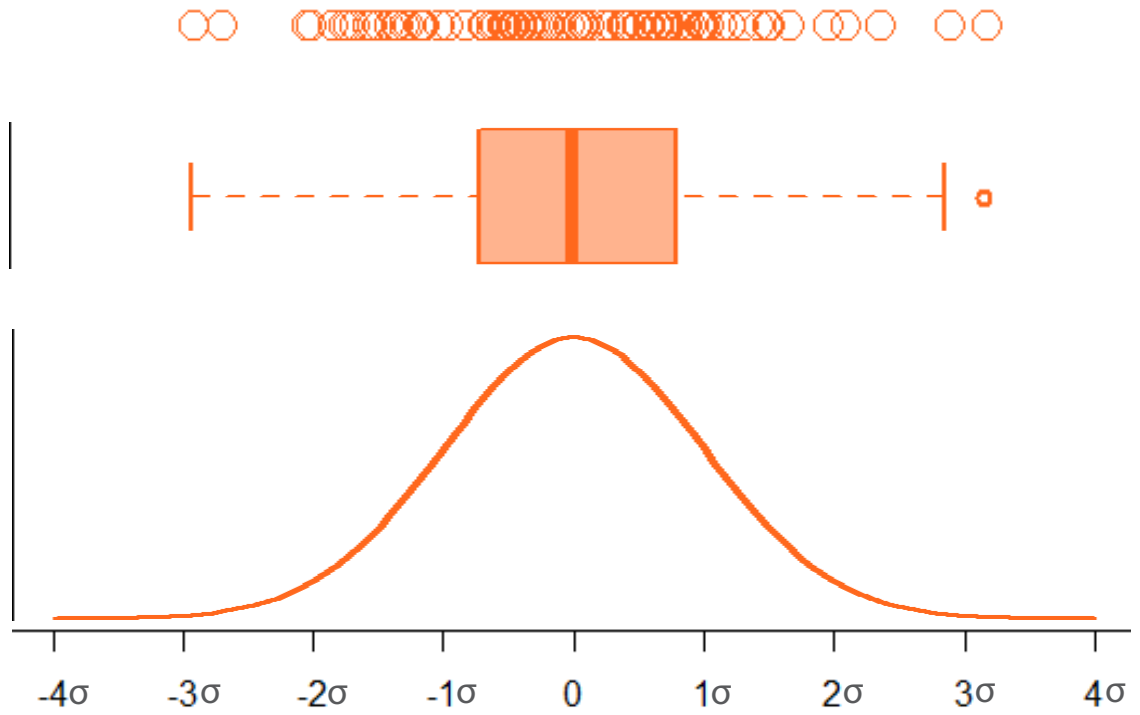


Quantitative Univariate Analysis

Mean

Median

Mode



Quantitative Univariate Analysis

Minimum

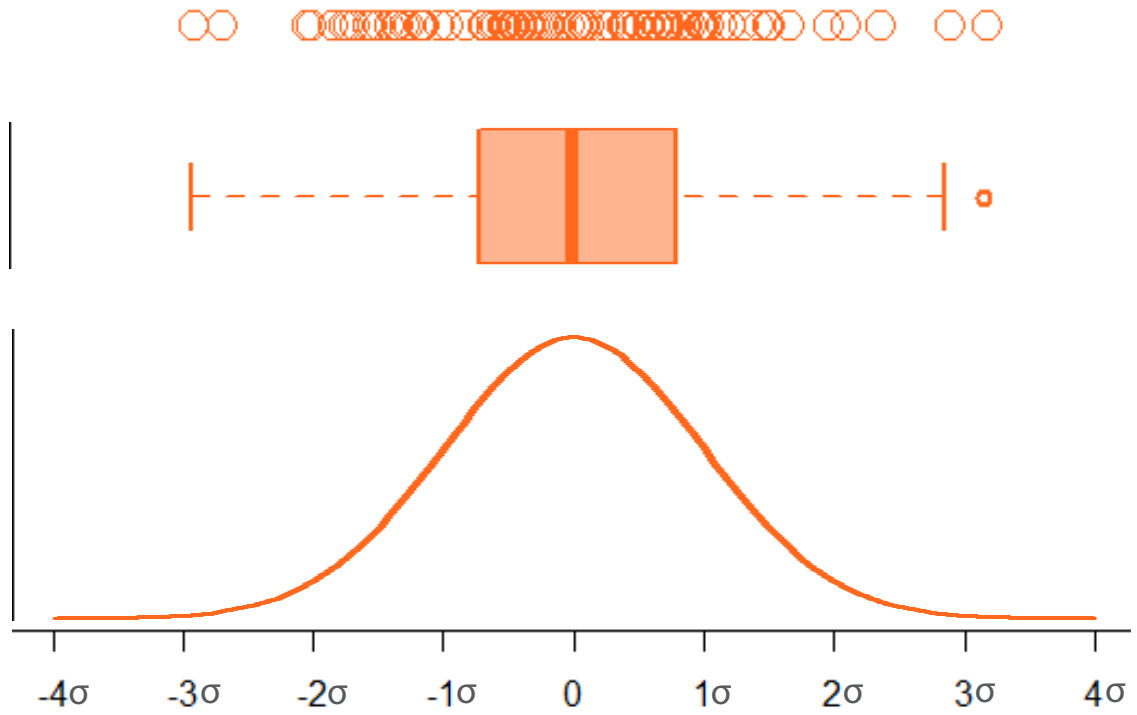
Maximum

Range

Quartiles

Variance

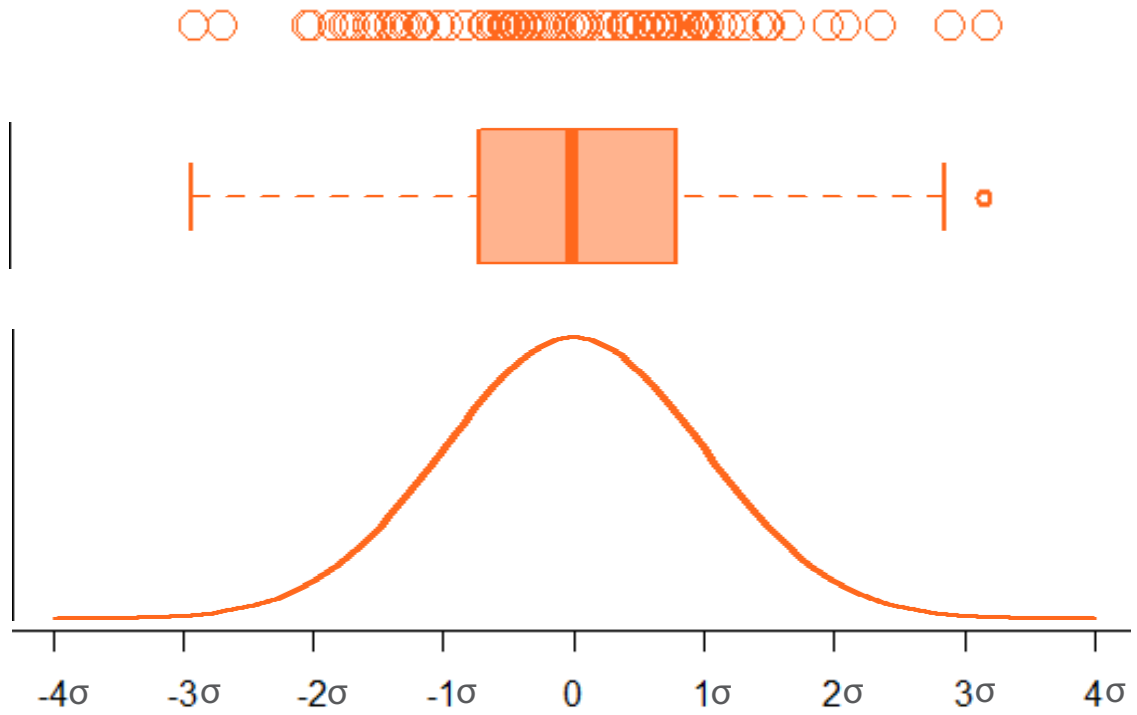
Standard deviation



Quantitative Univariate Analysis

Skewness

Kurtosis



Qualitative Bivariate Analysis

Joint frequency

Joint percentage

Marginal frequency

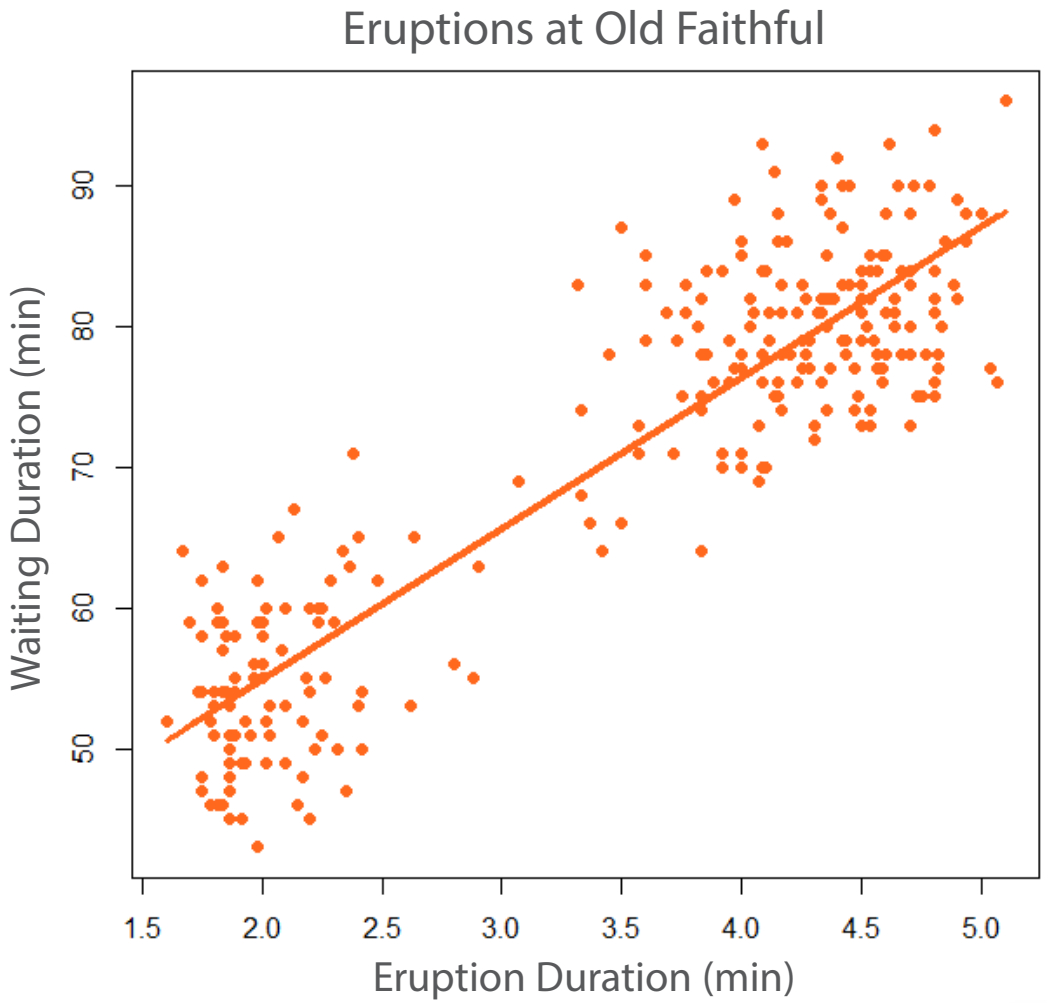
Movies by Genre and Rating					
Genre	G	PG	PG-13	R	Total
Action	2	70	311	229	612
Adventure	44	179	209	64	496
Animation	43	111	8	6	168
Comedy	45	258	472	506	1218
Drama	12	136	586	836	1570
Family	38	181	10	1	230
…	…	…	…	…	…
Total	230	1207	2686	3058	7181

Quantitative Bivariate Analysis

Relationship

Covariance

Correlation



Qualitative and Quantitative Bivariate Analysis

Quantitative measures

Partitioned by category

Average Critic Score by Genre	
Genre	Avg. Critic Score
Action	42.8
Adventure	51.3
Animation	56.1
Comedy	45.5
Drama	53.2
Horror	36.2
…	…

Guidance

Clean data

Domain knowledge

Understand biases

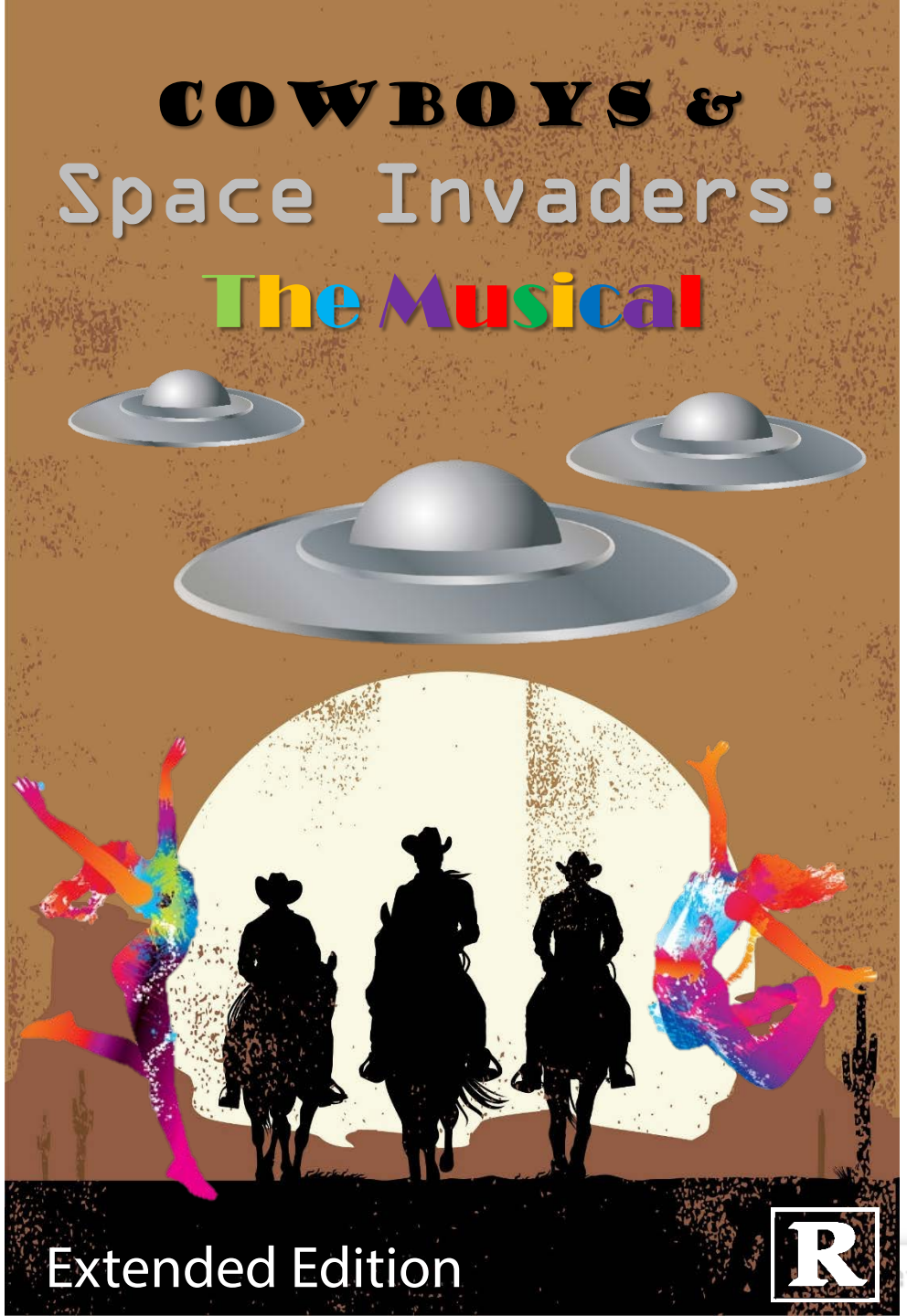
Guidance

Reproducibility

Understand implications

Know your limitations

Do Not Practice Advanced Data Analysis without the Proper Training!


COWBOYS &
Space Invaders:
The Musical
Extended Edition
R

Loading Data

Inspect Data

Univariate Qualitative Analysis

Univariate Quantitative Analysis (Location)

Univariate Quantitative Analysis (Spread)

Univariate Quantitative Analysis (Shape)

Univariate Quantitative Analysis (Summary)

Bivariate Qualitative Analysis

Bivariate Quantitative Analysis

Bivariate Qualitative and Quantitative Analysis

Summarize Table


COWBOYS &
Space Invaders:
The Musical
Extended Edition
R


COWBOYS &
Extended Edition
R


COWBOYS &
Space Invaders:
The Musical
Extended Edition

Summary

Introduction

Types of Analysis

Guidance

Demo