



Stages of Data Exploration



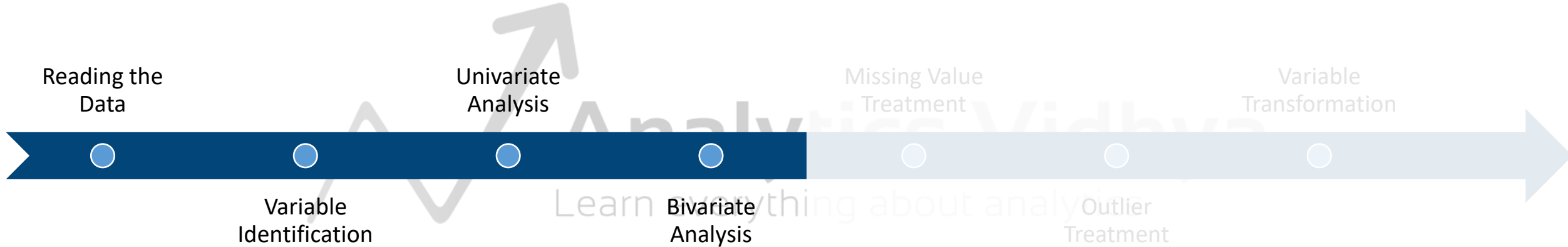
Stages of Data Exploration



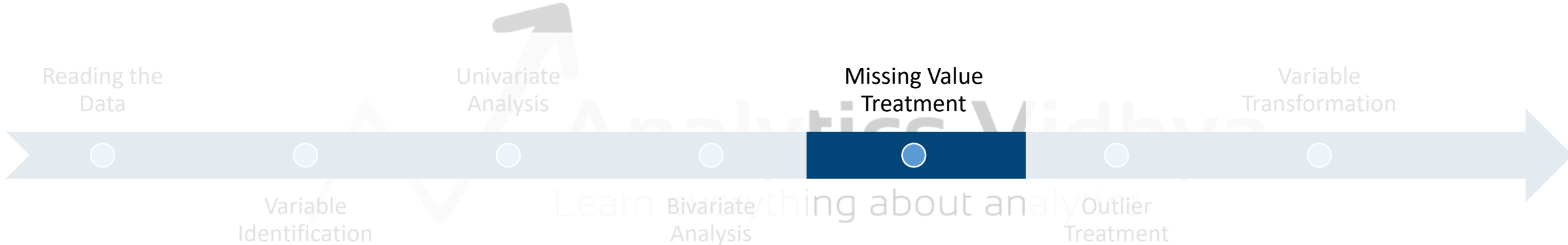
Stages of Data Exploration



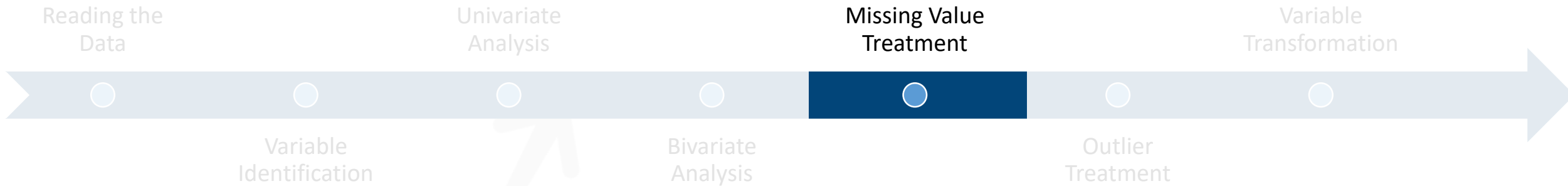
Stages of Data Exploration



Stages of Data Exploration



Stages of Data Exploration



Topics to be covered

1. Reasons for Missing Values?
2. Types of Missing Values?
3. Identifying Missing Values?
4. Different methods to deal with Missing Values

Analytics Vidhya
Learn everything about analytics

Missing Value Treatment

Reasons for Missing Values

1. Non-response – Example when you collect data on people's income and many choose not to answer.
2. Error in Data Collection
3. Error in Reading Data

Analytics Vidhya
Learn everything about analytics

Missing Value Treatment

Types of Missing Values

1. Missing Completely at Random (MCAR)
2. Missing at Random (MAR)
3. Missing not at Random (MNAR)

Analytics Vidhya
Learn everything about analytics

Types of Missing Values

Missing Completely at Random (MCAR)

Missing Values have no relation to –

1. The variable in which missing values exist.
2. Other Variables in the dataset.

Complete Data

Age	IQ
10	77
15	114
25	98
35	72
40	119
45	76
55	98
60	114
65	76
70	78
75	88
80	112
85	89
90	97

Incomplete Data

Age	IQ
10	
15	114
25	
35	72
40	
45	76
55	98
60	114
65	76
70	
75	88
80	112
85	
90	

Types of Missing Values

Missing at Random (MAR)

Missing Values have no relation to –

1. The variable in which missing values exist.

Missing Values have relation to –

1. The variables other than the variable in which missing values exist.

Complete Data

Age	IQ
10	77
15	114
25	98
35	72
40	119
45	76
55	98
60	114
65	76
70	78
75	88
80	112
85	89
90	97

Incomplete Data

Age	IQ
10	
15	
25	
35	
40	
45	
55	
60	114
65	76
70	78
75	88
80	112
85	89
90	97

Types of Missing Values

Missing not at Random (MNAR)

Missing Values have relation to –

1. The variable in which missing values exist.

Complete Data

Age	IQ
10	77
15	114
25	98
35	72
40	119
45	76
55	98
60	114
65	76
70	78
75	88
80	112
85	89
90	97

Incomplete Data

Age	MNAR
10	
15	114
25	98
35	
40	119
45	
55	98
60	114
65	
70	
75	
80	112
85	
90	97

Missing Value Treatment

Identifying Missing Values

describe()

Continuous Variables

isnull()

All Variables

Analytics Vidhya
Learn everything about analytics

Missing Value Treatment

Different methods to deal with missing values

1. Imputation
2. Deletion



Different methods to deal with Missing Values

Imputation



Different methods to deal with Missing Values

Imputation



Different methods to deal with Missing Values

Deletion

Age	Variable 1	Variable 2	Variable 3	Variable 4
10	77		47	
15	114	114	80	114
25	98		56	98
35	72	72	45	
40	119		74	119
45	76	76	56	
55	98	98	78	98
60	114	114	114	114
65	76	76	76	
70	78		78	
75	88	88	88	
80	112	112	112	112
85	89		89	
90	97		97	97

Row wise deletion

Age	Variable 1	Variable 2	Variable 3	Variable 4
10	77		47	
15	114	114	80	114
25	98		56	98
35	72	72	45	
40	119		74	119
45	76	76	56	
55	98	98	78	98
60	114	114	114	114
65	76	76	76	
70	78		78	
75	88	88	88	
80	112	112	112	112
85	89		89	
90	97		97	97

Column wise deletion

