Data Appendix

Overview: Four datasets were analyzed for this project, the datasets contained in the data files named RickAndMortyScripts.xls, Bing.xls, Afinn.xls, and NRC.xls. Therefore, this data appendix has four sections.

Section 1: RickAndMortyScript.xls

The unit of observation for this dataset – that is, what each row in the dataset represents – is a single line of dialogue spoken in the show *Rick and Morty*. Each line of dialogue is accompanied by additional information, such as in which episode the line is spoken. The variables included in the dataset are as follows:

Variable	Description	Example Responses
Season no.	The season number from which the line of dialogue originates	1, 2, 3
Episode no.	The episode number from which the line of dialogue originates	1, 7, 10
Episode name	The name of the episode from which the line of dialogue originates	"The Wedding Squanchers", "Tales from the Citadel"
Name	The name of the character who speaks the line	"Rick", "Morty", "Beth"
Line	The actual line of dialogue	"Is he going to die?", "Jerry, get a job."

The dataset includes 1905 rows (lines of dialogue) and contains no null values.

Section 2: Afinn.xls

The unit of observation for this dataset – that is, what each row in the dataset represents – is a word and a numeric classification of its emotional association The variables included in the dataset are as follows:

Variable	Description	Example Responses
Word	A plain English word	abnormal, silly, beg
Value	An integer between -5 and 5	-5, 0, 3

The dataset includes 2477 rows and contains no null values.

Section 3: Bing.xls

The unit of observation for this dataset – that is, what each row in the dataset represents – is a word and whether the word is associated with positive or negative feelings. The variables included in the dataset are as follows:

Variable	Description	Example Responses
Word	A plain English word	abnormal, silly, beg
Sentiment	Sentiment associated with the word	positive or negative

The dataset includes 6783 rows and contains no null values.

Section 4: NRC.xls

The unit of observation for this dataset – that is, what each row in the dataset represents – is a word and an emotion associated with that word. The variables included in the dataset are as follows:

Variable	Description	Example Responses
Word	A plain English word	abnormal, silly, beg
Sentiment	Sentiment associated with the word	joy, sadness, trust

The dataset includes 6468 rows and contains no null values.