

Rubric for Big Data:

	Mastery 40 to > 37 points	Approaching Mastery 36 to > 32 points	Progressing 31 to > 28 points	Emerging 27 to > 0 points	Incomplete
Deliverable 1: Perform ETL on Amazon Product Reviews	<ul style="list-style-type: none"> ✓ An Amazon Review dataset is extracted as a DataFrame (10 pt) ✓ The extracted dataset is transformed into four DataFrames with the correct columns (20 pt) ✓ ALL FOUR DataFrames are loaded into their respective tables in pgAdmin (10 pt) 	<ul style="list-style-type: none"> ✓ An Amazon Review dataset is extracted as a DataFrame (10 pt) ✓ The extracted dataset is transformed into four DataFrames with the correct columns, AND ONE df has incorrect column names (18 pt) ✓ THREE of the FOUR DataFrames are loaded into pgAdmin tables (8 pt) 	<ul style="list-style-type: none"> ✓ An Amazon Review dataset is extracted as a DataFrame (10 pt) ✓ The extracted dataset is transformed into four DataFrames, but the products df has duplicate rows, OR ONE or more dfs has incorrect column names (14 pt) ✓ TWO of the FOUR DataFrames are loaded into pgAdmin tables (7 pt) 	<ul style="list-style-type: none"> ✓ An Amazon Review dataset is extracted as a DataFrame (10 pt) ✓ The extracted dataset is transformed into four DataFrames, but the products df has duplicate rows, AND ONE or more dfs has incorrect column names (10 pt) ✓ TWO of the FOUR DataFrames are loaded into pgAdmin tables (7 pt) 	
	Mastery 40 to > 36 points	Approaching Mastery 36 to > 34 points	Progressing 34 to > 31 points	Emerging 31 to > 0 points	
Deliverable 2: Determine Bias of Vine Reviews	<ul style="list-style-type: none"> ✓ There is a DataFrame or table for the vine_table data (5 pt) ✓ The data is filtered to create a DataFrame or table where there are 20 or more total votes (5 pt) ✓ The data is filtered to create a DataFrame or table where the percentage of helpful_votes is equal to or greater than 50% (5 pt) ✓ The data is filtered to create a DataFrame or table where there is a Vine review (5 pt) ✓ The data is filtered to create a DataFrame or table where there isn't a Vine review (5 pt) ✓ The total number of votes, the number of 5 star reviews, and the percentage of 5 star reviews are calculated (15 pt) 	<ul style="list-style-type: none"> ✓ There is a DataFrame or table for the vine_table data (5 pt) ✓ The data is filtered to create a DataFrame or table where there are 20 or more total votes (5 pt) ✓ The data is filtered to create a DataFrame or table where the percentage of helpful_votes is equal to or greater than 50% (5 pt) ✓ The data is filtered to create a DataFrame or table where there is a Vine review (5 pt) ✓ The data is filtered to create a DataFrame or table where there isn't a Vine review (5 pt) ✓ The total number of votes and the number of 5 star reviews are calculated, but there is an error calculating the percentage of 5 star reviews (11 pt) 	<ul style="list-style-type: none"> ✓ There is a DataFrame or table for the vine_table data (5 pt) ✓ The data is filtered to create a DataFrame or table where there are 20 or more total votes (5 pt) ✓ The data is filtered to create a DataFrame or table where the percentage of helpful_votes is equal to or greater than 50% (5 pt) ✓ The data is filtered to create a DataFrame or table where there is a Vine review (5 pt) ✓ The data is filtered to create a DataFrame or table where there isn't a Vine review (5 pt) ✓ The total number of votes AND the number of 5 star reviews are calculated (9 pt) 	<ul style="list-style-type: none"> ✓ There is a DataFrame or table for the vine_table data (5 pt) ✓ The data is filtered to create a DataFrame or table where there are 20 or more total votes (5 pt) ✓ The data is filtered to create a DataFrame or table where the percentage of helpful_votes is equal to or greater than 50% (5 pt) ✓ The data is filtered to create a DataFrame or table where there is a Vine review (5 pt) ✓ The data is filtered to create a DataFrame or table where there isn't a Vine review (5 pt) ✓ Either the total number of votes OR the number of 5 star reviews are calculated (6 pt) 	<p>No submission was received</p> <p>-OR-</p> <p>Submission was empty or blank</p> <p>-OR-</p> <p>Submission contains evidence of academic dishonesty</p>

	Mastery 6 points to > 5 points	Approaching Mastery 5 to > 4 points	Progressing 4 to > 3 points	Emerging 3 to > 0 points	
Deliverable 3: Structure, Organization, and Formatting	<p>The written analysis has ALL of the following:</p> <ul style="list-style-type: none"> ✓ There is a title, and there are multiple sections. (2 pt) ✓ Each section has a heading and subheading. (2 pt) ✓ There are images and references to code, and they are formatted and displayed correctly. (2 pt) 	<p>The written analysis has ALL of the following:</p> <ul style="list-style-type: none"> ✓ There is a title, and there are multiple sections. (2 pt) ✓ Each section has a heading and subheading. (2 pt) ✓ There are images and references to code, and they are formatted and displayed correctly with one or two minor errors. (1 pt) 	<p>The written analysis has ALL of the following:</p> <ul style="list-style-type: none"> ✓ There is a title, and there are multiple sections. (2 pt) <p>AND ONE of the following:</p> <ul style="list-style-type: none"> ✓ Each section may have a heading and subheading. (2 pt) ✓ There are images and references to code, and they are formatted and displayed correctly with one or two minor errors. (1 pt) 	<p>The written analysis has ALL of the following:</p> <ul style="list-style-type: none"> ✓ There is a title. (1 pt) ✓ There may be a subheading for a section. (1 pt) ✓ There are no headings for each section, but there are three sections. (1 pt) 	
	Mastery 14 to > 12 points	Approaching Mastery 12 to > 10 points	Progressing 10 to > 8 points	Emerging 8 to > 0 points	
Deliverable 3: Analysis	<ul style="list-style-type: none"> ✓ The purpose is well defined (3 pt). ✓ ALL THREE questions are addressed in the results (7 pt) ✓ The summary states whether or not there is bias and is supported by the results, and there is a supporting additional recommendation (4 pt). 	<ul style="list-style-type: none"> ✓ The purpose is well defined (3 pt). ✓ ALL THREE questions are addressed in the results, but there is a misinterpretation (6 pt) ✓ The summary states whether or not there is bias and is supported by the results, but the additional recommendation doesn't support the statement (3 pt) 	<ul style="list-style-type: none"> ✓ The purpose is well defined (3 pt). ✓ TWO of the THREE questions are addressed in the results (5 pt) ✓ The summary states whether or not there is bias and is supported by the results, but there is no additional recommendation. (2 pt) 	<ul style="list-style-type: none"> ✓ The purpose is well defined (3 pt). ✓ ONE of the THREE questions is addressed in the results (4 pt). ✓ The summary states whether or not there is bias, but it is not supported by the results. (1 pt) 	