

# PARSE PRACTICE QUIZ

This practice quiz is designed to help students assess their understanding of the Acquire stage.

Parse Q1	Best Data Practices
Question Type	Multiple Choice
Topic	Best Practices
Question	You should ALWAYS
Correct Answer	Response Options
	Make a copy of your dataset before doing any data preparation.
	Make a copy of your dataset before cleaning the data in case a change you make is incorrect and you need to refer to the original.
X	All of the above.
	None of the above.
Grading Method	All or nothing
Points	1

Parse Q2	Best Data Practices
Question Type	True/False
Topic	Data Cleaning
Question	One of the most work-intensive tasks in cleaning data is handling inconsistent information.
Correct Answer	Answer Options
X	True
	False
Grading Method	Points earned for correct answer only.
Points	1

Parse Q3	Best Data Practices
Question Type	True/False
Topic	Datasets
Question	Most datasets are never 100% clean.
Correct Answer	Answer Options
X	True
	False
Grading Method	Points earned for correct answer only.
Points	1

Parse Q4	Best Data Practices
Question Type	Multiple Selection (M-S)
Topic	Data Processing
Question	After cleaning (mining) data, you should: <b>Check all that apply.</b>
Correct Answer	
X	Document everything you did to clean the data.
X	Be careful to write a set of instructions as you go, documenting exactly what is done in each step to identify bad data and which data points were removed.
	Filter out the data that you need and delete the data you don't need.
Grading Method	Right Minus Wrong Selection (+ - 0.2 pts per answer)
Points	1

Parse Q5	Best Data Practices																				
Question Type	Matching																				
Topic	Basic Statistical Methods																				
Question	Match the descriptions of the basic statistical methods.																				
Correct Answer	<table border="1"> <thead> <tr> <th>Term</th> <th></th> <th>Numeric Option</th> <th>Definition</th> </tr> </thead> <tbody> <tr> <td>Mean</td> <td></td> <td>1</td> <td>The most common method for finding averages.</td> </tr> <tr> <td>Median</td> <td></td> <td>2</td> <td>The most frequent response value.</td> </tr> <tr> <td>Mode</td> <td></td> <td>3</td> <td>The value in the precise center of the dataset.</td> </tr> <tr> <td></td> <td></td> <td>4</td> <td>Used to determine how far apart the most extreme values in a dataset are.</td> </tr> </tbody> </table>	Term		Numeric Option	Definition	Mean		1	The most common method for finding averages.	Median		2	The most frequent response value.	Mode		3	The value in the precise center of the dataset.			4	Used to determine how far apart the most extreme values in a dataset are.
Term		Numeric Option	Definition																		
Mean		1	The most common method for finding averages.																		
Median		2	The most frequent response value.																		
Mode		3	The value in the precise center of the dataset.																		
		4	Used to determine how far apart the most extreme values in a dataset are.																		
Grading Method	.333 points for each correct answer																				
Points	1																				