RMS® CCRA® Training Program
Exposure Data Quality Course Exercise

Learning Objectives:

Identify data quality resolution and completeness issues for a commercial California earthquake book of business. Where possible, also identify potentially suspect data values that indicate other data accuracy categories that may be compromised and list those categories.

Introduction:

The year is 2018. Your company's California commercial book of business was affected by an M 7.1 earthquake on the Puente Hills Fault in Southern California on September 7, 2017. Your modeled losses based on both event scenario and footprint analyses are lower than actual. RMS has performed a month-long claims analysis for this book of business and summarized their conclusions in a report.

The following files are available to you in order to answer the subsequent questions:

- Claims Analysis Report.pdf – Company A Claims Analysis Report from RMS
- Query Results CAEQ.xls – Excel spreadsheet containing exposure data profiles on 14 separate worksheets

PART I – Exposure Data Value Profiling

1. Listed below are the three "fitness for purpose" dependencies for any catastrophe exposed data analysis. For each category, use the information in the Excel spreadsheet and the Claims Analysis Report to provide data profiles, relative total insured values (e.g. percentages), and/or limits as metrics. One example is already provided for the peril category.

Peril	100% of the TIV (total insured values) is exposed to the earthquake peril.
Line of Business	
Hazard Region*	

**Hint: Refer to TIV by County tab of spreadsheet provided.*

2. Referring to the information in your answer to Question 1, and the data provided in the claims analysis report, would you focus your data quality efforts on just 2 or 3 primary data categories? If so, list which ones, and explain why. If not, explain why.

3. For each of the data categories listed below provide <u>one</u> data resolution or completeness issue and <u>one</u> data accuracy issue, and include a metric for each. One example is provided for you in the geocoding category.

Data Characteristics	Data Resolution or Completeness	Data Accuracy
Geocoding	*~89% of TIV is geocoded to street (high) resolution.*	*~50% of low resolution geocoded TIV clearly have missing address information (no street number, address or ZIP Code) or have information coded in the wrong data column.*
Construction Scheme		
Occupancy Class		

Data Characteristics	Data Resolution or Completeness	Data Accuracy
Construction Class		
Year Built		
Number of Stories		
Secondary Characteristics		
Policy Coding		
Treaty Coding		

PART II – Processes Impacting Data Quality

4. Each of the below data accuracy categories were discussed in Unit 3 of the course document, and sample data accuracy questions were provided. Using this information from the course material, provide one example of each category from the Company A portfolio. Use the data accuracy observations in Question 3 to help guide your answers. Finally, rank them in what you believe to be their order of importance. ***Note: The ranking is purely subjective and will be part of the discussion during first day of the instructor-led portion of the course.***

Rank	Data Accuracy Categories	Example
	Bias	
	Vintage	
	Validity	
	Consistency	
	Interpretability	