

Data Quality Dimensions

CustomerAddress		What can we do better?
Correct Values	Accuracy: 100%	
Data Fields with Values	Completeness: 100% Total record = 3999	
Values Free from Contradiction	Consistency: No	@column "state" Update New South Wales (has 86 records) to NWS and Victory (has 939 records) to VIC to be consistence.
Values Up to Date	Currency: No (since 2017)	
Data Items with Value Meta-data	Relevancy: customer_id	
Data Containing Allowable Values	Validity: Yes	
Records that are Duplicated	Uniqueness: No duplicate rows	
Presentation: https://seansothey.github.io/KPMG/CustomerAddress.html		

CustomerDemographic		What can we do better?
Correct Values	Accuracy: 100% @column "gender" = U is legit? Please clarify "U"?	@column "gender" If "U" is considered as unknown type and counted; it is acceptable. If not, 88 records of "U" should be updated to Female or Male.
Data Fields with Values	Completeness: 70% Total record = 4000	Check missing records @column <ul style="list-style-type: none"> • "last_name" has 125 • "DOB" has 87 • "job_title" has 506 • "job_industry_cate" has 656 • "tenure" has 87 Total missing is 1461
Values Free from Contradiction	Consistency: No	@column "gender" Update F to Female and M to Male "U" ?
Values Up to Date	Currency: No (since 2017)	
Data Items with Value Meta-data	Relevancy: customer_id	
Data Containing Allowable Values	Validity: Yes	
Records that are Duplicated	Uniqueness: No duplicate rows	
Presentation: https://seansothey.github.io/KPMG/CustomerDemographic.html		

*****Note: Customers below has no address record**

customer_id	first_name	last_name
3	Arlin	Dearle
10	Fiorenze	Birdall
22	Deanne	Durtnell
23	Olav	Polak

Transactions (path 3 months)		What can we do better?
Correct Values	Accuracy: 100%	
Data Fields with Values	Completeness: 99% Total record = 4996	Check missing records @column <ul style="list-style-type: none"> • "online_order" has 86 • "brand" has 60 • "product_line" has 60 • "product_class" has 60 • "product_size" has 60 • "standart_cost" has 60 • "product_first_sold_date" has 60 Total missing is 446
Values Free from Contradiction	Consistency: Yes	
Values Up to Date	Currency: No (since 2017)	
Data Items with Value Meta-data	Relevancy: customer_id	
Data Containing Allowable Values	Validity: Yes	
Records that are Duplicated	Uniqueness: No duplicate rows	
Presentation: https://seansothey.github.io/KPMG/Transactions3Months.html		

Prepared by SOTHEY SEAN