**DATA QUALITY CHALLENGES  
COMPLETENESS:**﻿whether all the required data is available or not

* **﻿**﻿PASSENGER\_ID
* PASSENGER\_NAME
* TRAIN\_NUMBER
* STATION\_NAME
* ARRIVAL\_ROUTE,DEPARTURE ROUTE

**ACCURACY**:correct values of data

* TRAIN\_NUMBER
* MOBILE\_NUMBER
* STATION\_NAME
* PASSENGER\_ID

**CONFORMANCE/VALIDITY**:correct format of data

* MOBILE\_NO:10 digits only
* PASSENGER\_NAME:strings only
* GENDER:no digits

**TIMELINESS**:data latency

* DISTANCE
* TIME
* AVAILABLE\_CLASS

**PROVENANCE**:proof for data correctness is major challenge

* ADDRESS
* STATION\_NAME
* PASSENGER\_NAME
* ADDRESS