Data Dictionary - Column Information

Column Name	Column Description	Expected/Allowed Values	Field Limitations	Additional Notes
Name of the column exactly as it appears in the dataset.	*A brief, plain-language explanation of what the data in the column means.*	*Specifies if there is an expected range and/or format of possible values. For example, if the data type is Date & Time, this field will note whether the timestamp is MM/DD/YYYY or MM/YYYY. If the Column Name is ice cream, this field might note that values can be Chocolate, Vanilla or Strawberry.* *If relevant, this field specifies the unit of measurement of the data field, e.g. thousands, millions, $ value, miles, feet, year, etc.*	*Describes any unique characteristics or potential analytical limitations presented by this field, including:* *- the reasoning for any null, zero, or empty values in the data* *- if the data in the column was integrated from another dataset or organization* *- if the data covered includes a different time period* *- the source of the column and how the data in the column was generated.* *For example, information on how the data in this column was generated can include whether the data was self-reported directly by a person, system generated by a database or agency system, derived through analytical manipulation of other fields or records; or obtained from a different agency.*	*Provides any additional relevant information about the data in the column, including:* *- definitions of acronyms, special term or codes, or jargon that appears in the field values;* *- the meaning of confusing or non-intuitive values in the data;* *- how the information in this column relates to information in other columns;* *- other unique details about this column.*
VendorID	A code indicating the TPEP provider that provided the record.	1=Creative Mobile Technologies, LLC 2=VeriFone Inc.		
tpep_pickup_datetime	The date and time when the meter was engaged.			
tpep_dropoff_datetime	The date and time when the meter was disengaged.			
passenger_count	The number of passengers in the vehicle.			This is a driver-entered value.
trip_distance	The elapsed trip distance in miles reported by the taximeter.			
RatecodeID	The final rate code in effect at the end of the trip.	1= Standard rate 2=JFK 3=Newark 4=Nassau or Westchester 5=Negotiated fare 6=Group ride 99 = Null/unknown		
store_and_fwd_flag	This flag indicates whether the trip record was held in vehicle memory before sending to the vendor, aka "store and forward," because the vehicle did not have a connection to the server.	Y= store and forward trip N= not a store and forward trip		
PULocationID	TLC Taxi Zone in which the taximeter was engaged			
DOLocationID	TLC Taxi Zone in which the taximeter was disengaged			
payment_type	A numeric code signifying how the passenger paid for the trip.	0= Flex Fare trip 1= Credit card 2= Cash 3= No charge 4= Dispute 5= Unknown 6= Voided trip		
fare_amount	The time-and-distance fare calculated by the meter. For additional information on the following columns, see https://www.nyc.gov/site/tlc/passengers/taxi-fare.page			
extra	Miscellaneous extras and surcharges.			
mta_tax	Tax that is automatically triggered based on the metered rate in use.			
tip_amount	Tip amount – This field is automatically populated for credit card tips. Cash tips are not included.			
tolls_amount	Total amount of all tolls paid in trip.			
improvement_surcharge	Improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015.			
total_amount	The total amount charged to passengers. Does not include cash tips.			
congestion_surcharge	Total amount collected in trip for NYS congestion surcharge.			
airport_fee	For pick up only at LaGuardia and John F. Kennedy Airports.			