

## Data Dictionary, 2017 Yellow Taxi Trip Data

This project utilizes a sample of the 2017 Yellow Taxi Trip Data, a massive dataset published by New York City's Taxi & Limousine Commission (TLC) as part of the NYC Open Data program. The original, complete dataset contains over 113 million rows and is available on the NYC Open Data website:

[https://data.cityofnewyork.us/Transportation/2017-Yellow-Taxi-Trip-Data/biws-g3hs/about\\_data](https://data.cityofnewyork.us/Transportation/2017-Yellow-Taxi-Trip-Data/biws-g3hs/about_data)

The sample dataset contains:

**22,699 rows** – each row represents a different trip

**18 columns**

Column name	Description
ID	Trip identification number
VendorID	A code indicating the TPEP provider that provided the record.  <b>1= Creative Mobile Technologies, LLC;</b>  <b>2= VeriFone Inc.</b>
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.
PULocationID	TLC Taxi Zone in which the taximeter was engaged
DOLocationID	TLC Taxi Zone in which the taximeter was disengaged
RateCodeID	The final rate code in effect at the end of the trip.  <b>1= Standard rate</b>  <b>2=JFK</b>  <b>3=Newark</b>  <b>4=Nassau or Westchester</b>

Column name	Description
	<b>5=Negotiated fare</b>  <b>6=Group ride</b>
Store_and_fwd_flag	This flag indicates whether the trip record was held in vehicle memory before being sent to the vendor, aka “store and forward,” because the vehicle did not have a connection to the server.  <b>Y= store and forward trip</b>  <b>N= not a store and forward trip</b>
Payment_type	A numeric code signifying how the passenger paid for the trip.  <b>1= Credit card</b>  <b>2= Cash</b>  <b>3= No charge</b>  <b>4= Dispute</b>  <b>5= Unknown</b>  <b>6= Voided trip</b>
Fare_amount	The time-and-distance fare calculated by the meter.
Extra	Miscellaneous extras and surcharges. Currently, this only includes the \$0.50 and \$1 rush hour and overnight charges.
MTA_tax	\$0.50 MTA tax that is automatically triggered based on the metered rate in use.
Improvement_surcharge	\$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015.
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.
Total_amount	The total amount charged to passengers. Does not include cash tips.