

Field Name Description	VendorID	Data Type	PK/FK	DataFrame Data Type
Trip_id	A code generated to uniquely identify the trip	Serial	PK	IntegerType
VendorID	A code indicating the TPEP provider that provided the record. 1= Creative Mobile Technologies, LLC; 2= VeriFone Inc.	int	FK	ShortType
tpep_pickup_datetime	The date and time when the meter was engaged.	Timestamp		TimestampType
tpep_dropoff_datetime	The date and time when the meter was disengaged.	Timestamp		TimestampType
Passenger_count	The number of passengers in the vehicle. This is a driver-entered value.	int		ShortType
Trip_distance	The elapsed trip distance in miles reported by the taximeter.	numeric(7,2)		DecimalType
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	int	FK	ShortType
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	char(1)		BooleanType
PULocationID	TLC Taxi Zone in which the taximeter was engaged	int		ShortType
DOLocationID	TLC Taxi Zone in which the taximeter was disengaged	int		ShortType
Payment_type	A numeric code signifying how the passenger paid for the trip. 1= Credit card 2= Cash 3= No charge 4= Dispute 5= Unknown 6= Voided trip	int	FK	ShortType
Fare_amount	The time-and-distance fare calculated by the meter.	numeric(7,2)		DecimalType
Extra	Miscellaneous extras and surcharges. Currently, this only includes the \$0.50 and \$1 rush hour and overnight charges.	numeric(7,2)		DecimalType
MTA_tax	\$0.50 MTA tax that is automatically triggered based on the metered rate in use.	numeric(7,2)		DecimalType
Tip_amount	Tip amount – This field is automatically populated for credit card tips. Cash tips are not included.	numeric(7,2)		DecimalType
Tolls_amount	Total amount of all tolls paid in trip.	numeric(7,2)		DecimalType
Improvement_surcharge	\$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015.	numeric(7,2)		DecimalType
Total_amount	The total amount charged to passengers. Does not include cash tips.	numeric(7,2)		DecimalType
congestion_surcharge		numeric(7,2)		DecimalType

Parquet Data Type	CSV Header
INT(32, false)	
INT(16, false)	VendorID
TIMESTAMP(isAdjustedToUTC=false, precision=MILLIS)	tpep_pickup_datetime
TIMESTAMP(isAdjustedToUTC=false, precision=MILLIS)	tpep_dropoff_datetime
INT(16, false)	passenger_count
DecimalType	trip_distance
INT(16, false)	RatecodeID
BOOLEAN	store_and_fwd_flag
INT(16, false)	PULocationID
INT(16, false)	DOLocationID
INT(16, false)	payment_type
DecimalType	fare_amount
DecimalType	extra
DecimalType	mta_tax
DecimalType	tip_amount
DecimalType	tolls_amount
DecimalType	improvement_surcharge
DecimalType	total_amount
DecimalType	congestion_surcharge