STORED PROCEDURE

CREATE PROCEDURE dbo.process\_presentation

AS

INSERT INTO dbo.nyctaxi\_yellow

SELECT

CASE WHEN nty.VendorID= 1 THEN 'Creative Mobile Technologies'

WHEN nty.VendorID = 2 THEN 'VeriFone'

ELSE 'Unknown'

END AS vendor,

FORMAT(nty.tpep\_pickup\_datetime,'yyyy-MM-dd') AS tpep\_pickup\_datetime,

FORMAT(nty.tpep\_dropoff\_datetime,'yyyy-MM-dd') AS tpep\_dropoff\_datetime,

lu1.Borough as pu\_borough,

lu1.Zone as pu\_zone,

lu2.Borough as do\_borough,

lu2.Zone as do\_zone,

CASE nty.payment\_type

WHEN 1 THEN 'Credit Card'

WHEN 2 THEN 'Cash'

WHEN 3 THEN 'No Charge'

WHEN 4 THEN 'Dispute'

WHEN 6 THEN 'Voided Trip'

ELSE 'Unknown'

END as payment\_method,

nty.passenger\_count as passenger\_count,

nty.trip\_distance as trip\_distance,

nty.total\_amount as total\_amount

FROM staging.nyctaxi\_yellow nty

LEFT JOIN staging.taxi\_zone\_lookup lu1

ON nty.PULocationID = lu1.LocationID

LEFT JOIN staging.taxi\_zone\_lookup lu2

ON nty.DOLocationID = lu2.LocationID

—--------------------------------------------------------------------------------------------------------------------------------------------------

CREATE PROCEDURE staging.datacleaning\_proc

@end\_date DATETIME2,

@start\_date DATETIME2

AS

DELETE from staging.nyctaxi\_yellow WHERE tpep\_pickup\_datetime < @start\_date OR tpep\_pickup\_datetime > @end\_date;

—---------------------------------------------------------------------------------------------------------------------------------------------------

CREATE PROCEDURE metadata.insert\_staging\_metadata

@pipeline\_run\_id VARCHAR(255),

@table\_name VARCHAR(255),

@processed\_date DATETIME2

AS

INSERT INTO metadata.processing\_log(pipeline\_run\_id,table\_processed,rows\_processed,latest\_processed\_pickup,processed\_datetime)

SELECT

@pipeline\_run\_id AS pipeline\_run\_id,

@table\_name AS table\_processed,

COUNT(\*) as rows\_processed,

MAX(tpep\_pickup\_datetime) AS latest\_processed\_picup,

@processed\_date AS processed\_datetime

FROM staging.nyctaxi\_yellow;

—---------------------------------------------------------------------------------------------------------------------------------------------------

CREATE PROCEDURE metadata.insert\_presentation\_metadata

@pipeline\_run\_id VARCHAR(255),

@table\_name VARCHAR(255),

@processed\_date DATETIME2

AS

INSERT INTO metadata.processing\_log(pipeline\_run\_id,table\_processed,rows\_processed,latest\_processed\_pickup,

processed\_datetime

)

SELECT @pipeline\_run\_id AS pipeline\_run\_id,

@table\_name as table\_processed,

COUNT(\*) AS rows\_processed,

MAX(tpep\_pickup\_datetime) AS latest\_processed\_pickup,

@processed\_date AS processed\_datetime

FROM dbo.nyctaxi\_yellow;

—--------------------------------------------------------------------------------------------------------------------------------------------------

CREATE TABLE

SELECT

'placeholder' AS pipeline\_run\_id,

'staging\_nyctaxi\_yellow' AS table\_processed,

COUNT(\*) as rows\_processed,

MAX(tpep\_pickup\_datetime) AS latest\_processed\_picup,

CURRENT\_TIMESTAMP AS processed\_datetime

FROM staging.nyctaxi\_yellow;

CREATE SCHEMA metadata;

CREATE TABLE metadata.processing\_log(

pipeline\_run\_id VARCHAR(255),

table\_processed VARCHAR(255),

rows\_processed INT,

latest\_processed\_pickup DATETIME2(6),

processed\_datetime DATETIME2(6)

);

—---------------------------------------------------------------------------------------------------------------------------------------------------

CREATE TABLE dbo.nyctaxi\_yellow

(

vendor VARCHAR(50),

tpep\_pickup\_datetime DATE,

tped\_dropoff\_datetime DATE,

pu\_borough VARCHAR(100),

pu\_zone VARCHAR(100),

do\_borough VARCHAR(100),

do\_zone VARCHAR(100),

payment\_method VARCHAR(50),

passenger\_count INT,

trip\_distance FLOAT,

total\_amount FLOAT

);

SELECT \* FROM dbo.nyctaxi\_yellow

—---------------------------------------------------------------------------------------------------------------------------------------------------