#################Difficulty uploading datasets

**- Executing (Error)**

**Messages**

* Error 0xc02020a1: Data Flow Task 1: Data conversion failed. The data conversion for column "end\_station\_name" returned status value 4 and status text "Text was truncated or one or more characters had no match in the target code page.".  
   (SQL Server Import and Export Wizard)
* Error 0xc020902a: Data Flow Task 1: The "Source - 202102-divvy-tripdata\_csv.Outputs[Flat File Source Output].Columns[end\_station\_name]" failed because truncation occurred, and the truncation row disposition on "Source - 202102-divvy-tripdata\_csv.Outputs[Flat File Source Output].Columns[end\_station\_name]" specifies failure on truncation. A truncation error occurred on the specified object of the specified component.  
   (SQL Server Import and Export Wizard)
* Error 0xc0202092: Data Flow Task 1: An error occurred while processing file "C:\Users\Christian\Documents\ExtractedDivy\202102-divvy-tripdata.csv" on data row 273.  
   (SQL Server Import and Export Wizard)

Error 0xc0047038: Data Flow Task 1: SSIS Error Code DTS\_E\_PRIMEOUTPUTFAILED. The PrimeOutput method on Source - 202102-divvy-tripdata\_csv returned error code 0xC0202092. The component returned a failure code when the pipeline engine called PrimeOutput(). The meaning of the failure code is defined by the component, but the error is fatal and the pipeline stopped executing. There may be error messages posted before this with more information about the failure.  
 (SQL Server Import and Export Wizard)

####Issue was resolved by converting some data sets to excel files and converting some data sets to access files then loading to sql server.

#####Difficulty executing a ‘union all’ of all data sets. Several data sets needed to have a column data type adjusted to nvarchar(255) for union to be successful.

ALTER TABLE [Divvy].[dbo].['202101-divvy-tripdata$']

ALTER COLUMN start\_station\_id NVARCHAR(255)

ALTER TABLE [Divvy].[dbo].['202102-divvy-tripdata$']

ALTER COLUMN start\_station\_id NVARCHAR(255)

ALTER TABLE [Divvy].[dbo].['202103-divvy-tripdata$']

ALTER COLUMN start\_station\_id NVARCHAR(255)

ALTER TABLE [Divvy].[dbo].['202104-divvy-tripdata$']

ALTER COLUMN end\_station\_id NVARCHAR(255)

ALTER TABLE [Divvy].[dbo].['202107-divvy-tripdata$']

ALTER COLUMN end\_station\_id NVARCHAR(255)

ALTER TABLE [Divvy].[dbo].[202111-divy-tripdata]

ALTER COLUMN start\_station\_id NVARCHAR(255)

ALTER TABLE [Divvy].[dbo].['202204-divvy-tripdata$']

ALTER COLUMN start\_station\_id NVARCHAR(255)

ALTER TABLE [Divvy].[dbo].['202207-divvy-tripdata$']

ALTER COLUMN start\_station\_id NVARCHAR(255)

ALTER TABLE [Divvy].[dbo].['202209-divvy-publictripdata$']

ALTER COLUMN end\_station\_id NVARCHAR(255)

ALTER TABLE [Divvy].[dbo].['202210-divvy-tripdata$']

ALTER COLUMN start\_station\_id NVARCHAR(255)

ALTER TABLE [Divvy].[dbo].['202211-divvy-tripdata$']

ALTER COLUMN start\_station\_id NVARCHAR(255)

ALTER TABLE [Divvy].[dbo].['202211-divvy-tripdata$']

ALTER COLUMN end\_station\_id NVARCHAR(255)

ALTER TABLE [Divvy].[dbo].['202212-divvy-tripdata$']

ALTER COLUMN end\_station\_id NVARCHAR(255)

SELECT \*

INTO [Divvy].[dbo].[UnionDivvy]

FROM [Divvy].[dbo].[202101-divvy-tripdata]

UNION ALL

SELECT \*

FROM [Divvy].[dbo].[202102-divvy-tripdata]

UNION ALL

SELECT \*

FROM [Divvy].[dbo].[202103-divvy-tripdata]

UNION ALL

SELECT \*

FROM [Divvy].[dbo].[202104-divvy-tripdata]

UNION ALL

SELECT \*

FROM [Divvy].[dbo].[202105-divvy-tripdata]

UNION ALL

SELECT \*

FROM [Divvy].[dbo].[202106-divvy-tripdata]

UNION ALL

SELECT \*

FROM [Divvy].[dbo].[202107-divvy-tripdata]

UNION ALL

SELECT \*

FROM [Divvy].[dbo].[202108-divvy-tripdata]

UNION ALL

SELECT \*

FROM [Divvy].[dbo].[202109-divvy-tripdata]

UNION ALL

SELECT \*

FROM [Divvy].[dbo].[202110-divvy-tripdata]

UNION ALL

SELECT \*

FROM [Divvy].[dbo].[202111-divvy-tripdata]

UNION ALL

SELECT \*

FROM [Divvy].[dbo].[202112-divvy-tripdata]

UNION ALL

SELECT \*

FROM [Divvy].[dbo].['202201-divvy-tripdata$']

UNION ALL

SELECT \*

FROM [Divvy].[dbo].['202202-divvy-tripdata$']

UNION ALL

SELECT \*

FROM [Divvy].[dbo].['202203-divvy-tripdata$']

UNION ALL

SELECT \*

FROM [Divvy].[dbo].['202204-divvy-tripdata$']

UNION ALL

SELECT \*

FROM [Divvy].[dbo].['202205-divvy-tripdata$']

UNION ALL

SELECT \*

FROM [Divvy].[dbo].['202206-divvy-tripdata$']

UNION ALL

SELECT \*

FROM [Divvy].[dbo].['202207-divvy-tripdata$']

UNION ALL

SELECT \*

FROM [Divvy].[dbo].['202208-divvy-tripdata$']

UNION ALL

SELECT \*

FROM [Divvy].[dbo].['202209-divvy-publictripdata$']

UNION ALL

SELECT \*

FROM [Divvy].[dbo].['202210-divvy-tripdata$']

UNION ALL

SELECT \*

FROM [Divvy].[dbo].['202210-divvy-tripdata$']

UNION ALL

SELECT \*

FROM [Divvy].[dbo].['202211-divvy-tripdata$']

###Getting rid of duplicate rides and extracting date names, parts, and length of ride from started\_at column

SELECT DISTINCT [ride\_id]

,[rideable\_type]

,[started\_at]

,[ended\_at]

,[start\_station\_name]

,[start\_station\_id]

,[end\_station\_name]

,[end\_station\_id]

,[start\_lat]

,[start\_lng]

,[end\_lat]

,[end\_lng]

,[member\_casual]

,Datename(weekday FROM started\_at) AS Dayofweek

,Datepart(weekday FROM started\_at) AS DayofweekNum

,Datepart(month FROM started\_at) AS Month\_Num

,Datename(month FROM started\_at) AS Month

,Datepart(hour FROM started\_at) AS HourofDay

,Datepart(year FROM started\_at) AS Year

,DateDiff(minute, started\_at, ended\_at) AS ride\_length\_min

INTO Divvy.dbo.UnionDivvy2

FROM Divvy.dbo.UnionDivvy

###Selecting all data where ride length is 24 hours or greater—Ride is considered a stolen bike

SELECT \*

INTO Divvy2.dbo.Rides\_Over\_24hrs

FROM Divvy2.dbo.CleanDivvy

WHERE ride\_length\_min > 1440

SELECT DISTINCT start\_station\_name

,count(\*) AS Bikes\_stolen

INTO Divvy2.dbo.Num\_bikes\_stolen\_from\_station

FROM Divvy2.dbo.Rides\_Over\_24hrs

GROUP BY start\_station\_name

ORDER BY Bikes\_stolen DESC

SELECT DISTINCT [Dayofweek]

,Count(\*) AS StolenBikes

,

INTO Divvy2.dbo.Daily\_StolenBikes

FROM [Divvy2].[dbo].[Rides\_Over\_24hrs]

GROUP BY Dayofweek

ORDER BY StolenBikes DESC

SELECT rideable\_type

,count(\*) AS bikes\_stolen

INTO Divvy2.dbo.Num\_of\_type\_StolenBikes

FROM Divvy2.dbo.Rides\_Over\_24hrs

GROUP BY rideable\_type

SELECT DISTINCT HourofDay

,Count(\*) AS StolenBikes

INTO Divvy2.dbo.StolenBikes\_HourOfDay

FROM Divvy2.dbo.Rides\_Over\_24hrs

GROUP BY HourofDay

ORDER BY HourofDay ASC

SELECT DISTINCT Month

,Month\_Num

,Year

,Count(\*) AS StolenBikes

INTO Divvy2.dbo.Month\_Year\_StolenBikes

FROM Divvy2.dbo.Rides\_Over\_24hrs

GROUP BY Month

,month\_num

,Year

ORDER BY Month\_Num ASC