**#Module 20 Final Project**

**## Week1 – Data Source and ETL**

**## Tables**

* **Routes\_GarbageRecycle2015RR**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column name** | **What is it?** | **Unique** | **Comments - Actions** |
| LANDFILL | N/a | 1 | DROP? All have the same |
| the\_geom | Route Coordinates | 184 |  |
| GARB\_RT | Route names/ID | 184 | Key |
| GARB\_DAY | Day of the week when it runs | Mon - Fri |  |
| GARB\_SUP | Supervisor's name | 9 |  |
| SUPER\_NUM | Supervisor's ID | 6 | DROP? Match BARB\_SUP but some names do NOT have ID |
| OP\_TYPE | Vehicle type | 2 |  |
| RT\_OLD | Route old names | 91 | DROP? Unique but half are blank |

1\_Routes\_GarbageRecycle2015RR here (link here)

* **Waste\_Collection\_\_\_Diversion\_Report\_\_daily\_**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column name** | **What is it?** | **Unique** | **Comments - Actions** |
| Report Date | Date of the collection | No | from 1/12/2003 to 5/15/2021 Note that there are a couple of bad entries (year 30) |
| Load Type | Type of garbage | 21 | All those for which we have route coordinates ARE the same type: ‘GARBAGE COLLECTIONS’ |
| Load Time | Date with time in 24 hr HH:MM format (M/DD/YY HH:MM) | Yes | Even though is unique, it might be too complicated to use as unique key |
| Load Weight | Weight of the load | No | Need to find out which units (pounds, tons, etc) |
| Dropoff Site | Drop off collection site | 34 | To three sites collect 81% and top 7 99% |
| Route Type | Type of route | 41 | Garbage collection more than half, top five 90% top 9, 99% |
| Route Number | Route Number | 2021 | About 500 have NO load weight  We’ll be working with only the 184 that appear on the Routes\_GarbageRecycle |
| Load ID | Unique trip ID | 733252 | Unique numbers ranging between 101223 to 921299 |

2\_Waste\_Collection\_\_\_Diversion\_Report\_\_daily here (link here)

* **Routes\_with\_coordanates(export from Tableau)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column name** | **What is it?** | **Unique** | **Comments - Actions** |
| Garb Rt | Route names/ID from Routes\_GarbageRecycle2015RR | 184 |  |
| Geometry (Garbage Routes.geojson) | Route Coordinates (the geom column) from Routes\_GarbageRecycle2015RR |  |  |
| Latitude (generated) | Latitude coordinate with multiple decimal points |  |  |
| Longitude (generated) | Latitude coordinate with multiple decimal points |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |