```
# First step was to install openpyxl to make it possible to read excel files
!pip install openpyxl==3.1.2
```

```
Requirement already satisfied: openpyxl==3.1.2 in /root/venv/lib/python3.9/site-packages (3.1.2)
Requirement already satisfied: et-xmlfile in /root/venv/lib/python3.9/site-packages (from openpyxl==3.1.2) (1.1.0)

[notice] A new release of pip is available: 23.0.1 -> 24.0
[notice] To update, run: pip install --upgrade pip
```

```
# Import important libraries
import numpy as np
import pandas as pd
import openpyxl
```

```
df = pd.read_excel("UPDATED_BMMS_overview.xlsx")
```

```
df
```

	road object N1 3.9% N2 2.7% 706 others 93.4%	km float64 0.004 - 522.718	type object Box Culvert 44% RCC Girder ... 19.1% 12 others 36.9%	LRPName object LRP003a 2.3% LRP001a 2.2% 1492 others 95.4%	name object	length float64	condition object	structureNr int64
0	N1	1.8	Box Culvert	LRP001a	.	11.3	A	117861
1	N1	4.925	Box Culvert	LRP004b	.	6.6	A	117862
2	N1	8.976	PC Girder Bridge	LRP008b	Kanch pur Bridge.	394.23	A	119889
3	N1	10.88	Box Culvert	LRP010b	NOYAPARA CULV...	6.3	A	112531
4	N1	10.897	Box Culvert	LRP010c	ADUPUR CULVERT	6.3	A	112532
5	N1	11.296	Box Culvert	LRP011a	NAYABARI KASP...	8.3	A	101110
6	N1	12.239	Box Culvert	LRP012a	KHAS PARA BOX ...	9.3	A	101117
7	N1	12.253	Box Culvert	LRP012b	DAWAN BAG BOX...	6.1	A	101119
8	N1	12.66	PC Girder Bridge	LRP013a	Madanpur Bridge....	27.5	A	119897
9	N1	12.66	PC Girder Bridge	LRP013a	MADAN PUR (R)	26.3	A	109841

20415 rows, showing 10 per page Page 1 of 2042

```
# Filter on N1
df_n1 = df[df['road']=='N1']
df_n1
```

	road object N1 100%	km float64 1.8 - 460.113	type object Box Culvert 46.5% RCC Girder ... 25.8% 7 others 27.7%	LRPName object LRP031a 0.5% LRP013a 0.4% 637 others 99.1%	name object . 0.5% SOUTH MIT... 0.5% 744 others 99%	length float64 0.8 - 1408.8	condition object A 59.2% B 19.7% 2 others 21.1%	structureNr int64 100403 - 121361
0	N1	1.8	Box Culvert	LRP001a	.	11.3	A	117861
1	N1	4.925	Box Culvert	LRP004b	.	6.6	A	117862
2	N1	8.976	PC Girder Bridge	LRP008b	Kanch pur Bridge.	394.23	A	119889
3	N1	10.88	Box Culvert	LRP010b	NOYAPARA CULV...	6.3	A	112531
4	N1	10.897	Box Culvert	LRP010c	ADUPUR CULVERT	6.3	A	112532
5	N1	11.296	Box Culvert	LRP011a	NAYABARI KASP...	8.3	A	101110
6	N1	12.239	Box Culvert	LRP012a	KHAS PARA BOX ...	9.3	A	101117
7	N1	12.253	Box Culvert	LRP012b	DAWAN BAG BOX...	6.1	A	101119
8	N1	12.66	PC Girder Bridge	LRP013a	Madanpur Bridge....	27.5	A	119897
9	N1	12.66	PC Girder Bridge	LRP013a	MADAN PUR (R)	26.3	A	109841

787 rows, showing 10 per page Page 1 of 79

Since we only need to take the road from Dhaka to Chittagong, we decided to look in the java file and select a LRP in Chittagong. This is the LRP with LRPName LRP249. Then, we only selected the rows with LRP's above this LRP.

```
#Select road until LRP249a
df_n1 = df_n1.sort_values(by = 'LRPName')
index_of_LRP249 = df_n1.index[df_n1['LRPName'] == 'LRP249a'][0]
filtered_df = df_n1.loc[:index_of_LRP249]
```

```
filtered_df
```

	road object N1 100%	km float64 1.8 - 253.777	type object Box Culvert 53.7% PC Girder B... 23.6% 6 others 22.6%	LRPName object LRP031a 1.4% LRP013a 1% 237 others 97.6%	name object . 1.4% SHIKALBAH... 0.7% 289 others 98%	length float64 1.0 - 1408.8	condition object A 58.4% B 27.7% 2 others 13.9%	structureNr int64 100545 - 121361
0	N1	1.8	Box Culvert	LRP001a	.	11.3	A	117861
1	N1	4.925	Box Culvert	LRP004b	.	6.6	A	117862
2	N1	8.976	PC Girder Bridge	LRP008b	Kanch pur Bridge.	394.23	A	119889
127..	N1	8.976	PC Girder Bridge	LRP008b	KANCHPUR PC G...	397	C	101102
145..	N1	10.543	Box Culvert	LRP010a	KATCHPUR BOX ...	8	B	101106
3	N1	10.88	Box Culvert	LRP010b	NOYAPARA CULV...	6.3	A	112531
4	N1	10.897	Box Culvert	LRP010c	ADUPUR CULVERT	6.3	A	112532
5	N1	11.296	Box Culvert	LRP011a	NAYABARI KASP...	8.3	A	101110
145..	N1	11.808	Box Culvert	LRP011c	NAYABARI BOX C...	10.6	B	101114
6	N1	12.239	Box Culvert	LRP012a	KHAS PARA BOX ...	9.3	A	101117

296 rows, showing 10 per page Page 1 of 30

We decided to create a new column called LRP_filtered, where only the first 6 digits of the LRP are showed. Based on this row, we delete duplicate bridges if they are on the same LRP. The focus is only on the first 6 digits since there are bridges with the same name on for example LRPa and LRPb. When we would focus on the whole LRP, these would not get deleted

```
#Filter on the first 6 digits
filtered_df['LRP_filtered'] = filtered_df['LRPName'].str[:6]
```

```
/tmp/ipykernel_217/2857888424.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  filtered_df['LRP_filtered'] = filtered_df['LRPName'].str[:6]
```

```
filtered_df = filtered_df.drop_duplicates(subset=['LRP_filtered'])
```

```
filtered_df
```

	road object N1 100%	km float64 1.8 - 253.777	type object Box Culvert 67.7% PC Girder B... 12.2% 4 others 20.1%	LRPName object LRP001a 0.6% LRP004b 0.6% 162 others 98.8%	name object . 1.8% Kanch pur Br... 0.6% 160 others 97.6%	length float64 1.0 - 1408.8	condition object A 61% B 27.4% C 11.6%	structureNr int64 100545 - 121361
0	N1	1.8	Box Culvert	LRP001a	.	11.3	A	117861
1	N1	4.925	Box Culvert	LRP004b	.	6.6	A	117862
2	N1	8.976	PC Girder Bridge	LRP008b	Kanch pur Bridge.	394.23	A	119889
145..	N1	10.543	Box Culvert	LRP010a	KATCHPUR BOX ...	8	B	101106
5	N1	11.296	Box Culvert	LRP011a	NAYABARI KASP...	8.3	A	101110
6	N1	12.239	Box Culvert	LRP012a	KHAS PARA BOX ...	9.3	A	101117
9	N1	12.66	PC Girder Bridge	LRP013a	MADAN PUR (R)	26.3	A	109841
12	N1	13.574	Box Culvert	LRP014a	KAWTALA BOX C...	11.9	A	109794
145..	N1	15.465	Box Culvert	LRP016a	MALIBAG BOX C...	5.55	B	109800
145..	N1	17.134	PC Girder Bridge	LRP017b	LANGOLBANDO ...	159.5	B	109808

164 rows, showing 10 per page Page 1 of 17

```
#A new column is created where the type is shown, this is needed for the python files.
filtered_df['model_type'] = 'bridge'
```

```
/tmp/ipykernel_217/1403558431.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  filtered_df['model_type'] = 'bridge'
```

```
filtered_df
```

	road object	km float64	type object	LRPName object	name object (baro)	length float64	condition object	structureNr int64
95	N1	167.963	Box Culvert	LRP169a	BAROYARHAT BO...	3	A	101236
127...	N1	205.615	RCC Girder Bridge	LRP208a	BAROCONDO BRI...	22.25	C	100583
139	N1	206.866	Box Culvert	LRP209a	BARO-CUNDHO...	10.2	A	100586

3 rows, showing 10 per page Page 1 of 1

```
#Now we open the read the road dataset
roads = pd.read_csv('_roads3.csv')
roads
```

	road object (N1 2.6%, N5 2.4%, 845 others 95%)	chainage float64	lrp object	lat float64	lon float64	gap object	type object	name object
60	N1	22.629	LRP023	23.6173604	90.6096944	nan	KmPost	Km post missing
61	N1	23.564	LRP023a	23.6095556	90.6121944	nan	Bridge	Meghna bridge
62	N1	23.629	LRP024	23.609	90.6124722	nan	KmPost	Km post missing
63	N1	24.503	LRP024a	23.6018611	90.6159722	nan	Bridge	Bridge end
64	N1	24.629	LRP025	23.6009444	90.6165556	nan	KmPost	Km post missing
65	N1	25.629	LRP026	23.5959997	90.624666	nan	KmPost	Km post missing
66	N1	25.639	LRP026a	23.5959441	90.6246938	nan	Culvert	Box culvert
67	N1	26.313	LRP026b	23.5917497	90.6293327	nan	Culvert	Box culvert
68	N1	26.629	LRP027	23.5896108	90.6313882	nan	KmPost	Km post missing
69	N1	27.503	LRP027a	23.5837497	90.6369441	nan	Bridge	Bridge start

51348 rows, showing 10 per page Page 7 of 5135

```
#Only select the N1 road, since this is our focus
roads_n1 = roads[roads['road']=='N1']
roads_n1
```

	road object	chainage float64	lrp object (lrp249)	lat float64	lon float64	gap object	type object	name object
579	N1	245.456	LRP249	22.3314716	91.8515556	nan	KmPost	Km post missing
580	N1	245.665	LRP249a	22.3297493	91.8521111	nan	Bridge	Shah Amanath (K...

2 rows, showing 10 per page Page 1 of 1

```
# Here we do the same as in the bridges dataset, only the LRPs above LRP249a are selected
roads_n1 = roads_n1.sort_values(by = 'lrp')
roads_n1 = pd.concat([roads_n1.iloc[-3:], roads_n1.iloc[:-3]])
index_of_LRP249_roads = roads_n1.index[roads_n1['lrp'] == 'LRP249a'][0]
filtered_roads = roads_n1.loc[:index_of_LRP249_roads]
```

filtered_roads

	road object N1 100%	chainage float64 0.0 - 245.665	lrp object LRPS 0.2% LRPSa 0.2% 579 others 99.7%	lat float64 22.3297493 - 23.7...	lon float64 90.443333 - 91.85...	gap object BS 11.2% BE 11.2% Missing 77.6%	type object KmPost 41% Culvert 26.3% 11 others 32.7%	name object Km post mi... 27.2% Box culvert 25.3% 114 others 47.5%
0	N1	0	LRPS	23.7060278	90.443333	nan	Others	Start of Road afte...
1	N1	0.814	LRPSa	23.7029167	90.4504167	nan	Culvert	Box Culvert
2	N1	0.822	LRPSb	23.7027778	90.4504722	nan	CrossRoad	Intersection with ...
3	N1	1	LRP001	23.7021389	90.4519722	nan	KmPost	Km post missing
4	N1	2	LRP002	23.6978886	90.4605833	nan	KmPost	Km post missing
5	N1	2.13	LRP002a	23.6973608	90.4616667	nan	Culvert	Box culvert
6	N1	3	LRP003	23.693833	90.4691382	nan	KmPost	Km post missing
7	N1	4	LRP004	23.6936108	90.4787771	nan	KmPost	Km post missing
8	N1	4.175	LRP004a	23.6938052	90.4805271	nan	SideRoad,Right	Road to Narayan...
9	N1	5	LRP005	23.6947497	90.4884997	nan	KmPost	Km post missing

581 rows, showing 10 per page Page 1 of 59

```
# A new column is added, and the type of the roads is "link"
filtered_roads['model_type'] = 'link'
```

```
/tmp/ipykernel_217/3636629436.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  filtered_roads['model_type'] = 'link'
```

```
#To calculate the length column, we calculate the difference between chainage of the previous point

filtered_roads['length'] = filtered_roads['chainage'].shift(-1) - filtered_roads['chainage']
filtered_roads
```

```
/tmp/ipykernel_217/4180423981.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  filtered_roads['length'] = filtered_roads['chainage'].shift(-1) - filtered_roads['chainage']
```

	road object N1 100%	chainage float64 0.0 - 245.665	lrp object LRPS 0.2% LRPSa 0.2% 579 others 99.7%	lat float64 22.3297493 - 23.7...	lon float64 90.443333 - 91.85...	gap object BS 11.2% BE 11.2% Missing 77.6%	type object KmPost 41% Culvert 26.3% 11 others 32.7%	name object Km post mi... 27.2% Box culvert 25.3% 114 others 47.5%
0	N1	0	LRPS	23.7060278	90.443333	nan	Others	Start of Road afte...
1	N1	0.814	LRPSa	23.7029167	90.4504167	nan	Culvert	Box Culvert
2	N1	0.822	LRPSb	23.7027778	90.4504722	nan	CrossRoad	Intersection with ...
3	N1	1	LRP001	23.7021389	90.4519722	nan	KmPost	Km post missing
4	N1	2	LRP002	23.6978886	90.4605833	nan	KmPost	Km post missing
5	N1	2.13	LRP002a	23.6973608	90.4616667	nan	Culvert	Box culvert
6	N1	3	LRP003	23.693833	90.4691382	nan	KmPost	Km post missing
7	N1	4	LRP004	23.6936108	90.4787771	nan	KmPost	Km post missing
8	N1	4.175	LRP004a	23.6938052	90.4805271	nan	SideRoad,Right	Road to Narayan...
9	N1	5	LRP005	23.6947497	90.4884997	nan	KmPost	Km post missing

581 rows, showing 10 per page Page 1 of 59

filtered_df

	road object N1 100%	km float64 1.8 - 253.777	type object Box Culvert 67.7% PC Girder B... 12.2% 4 others 20.1%	LRPName object LRP001a 0.6% LRP004b 0.6% 162 others 98.8%	name object . 1.8% Kanch pur Br... 0.6% 160 others 97.6%	length float64 1.0 - 1408.8	condition object A 61% B 27.4% C 11.6%	structureNr int64 100545 - 121361
145..	N1	17.722	PC Girder Bridge	LRP018a	DARIKANDI BRID...	20	B	109853
145..	N1	18.742	PC Girder Bridge	LRP019a	MOLLIK PARA P.C...	40.5	B	109810
16	N1	19.76	Box Culvert	LRP020a	KRIBORDI BOX C...	6	A	109817
19	N1	21.184	PC Girder Bridge	LRP021c	MARIKHALL BRID...	89.2	A	101103
145..	N1	22.328	Box Culvert	LRP022a	SANGAKADI CUL...	5.5	B	101112
20	N1	22.596	PC Girder Bridge	LRP023a	ASHIR CHAR (Rig...	224.8	A	109820
24	N1	24.393	PC Girder Bridge	LRP024a	Meghna Bridge	924.85	A	119967
25	N1	26.115	Box Culvert	LRP026a	TATULTALA CUL...	5.6	A	101132
27	N1	27.12	Box Culvert	LRP027a	MALUAKANDI CU...	5.6	A	101150
29	N1	28.315	PC Girder Bridge	LRP028a	Bhater char Bridg...	173.25	A	119973

164 rows, showing 10 per page Page 2 of 17

```
#Since length of the chainage was in kilometers, we needed to multiply the length by 1000
filtered_roads['length'] = filtered_roads['length']*1000
```

```
/tmp/ipykernel_217/2475612705.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  filtered_roads['length'] = filtered_roads['length']*1000
```

Now it is time to merge both dataframes. First, some column names needed to be renamed. In addition, the three LRPs for the start are now at the end due to the sort_values, therefore, we take the last three rows put them on top

```
# Now it is time to merge both dataframes
filtered_df = filtered_df.rename(columns = {'LRPName': 'lrp'})
filtered_roads = filtered_roads.rename(columns = {'roadName':'name'} )
total_df = pd.concat([filtered_df, filtered_roads])
total_df = total_df.sort_values(by='lrp')

#Fix the starting point of the LRPs
total_df = pd.concat([total_df.iloc[-3:], total_df.iloc[:-3]])
total_df
```

	road object N1 100%	km float64 1.8 - 253.777	type object KmPost 31.9% Culvert 20.5% 17 others 47.5%	lrp object LRP008b 0.3% LRP010a 0.3% 622 others 99.5%	name object Km post mi... 21.2% Box culvert 19.7% 276 others 59.1%	length float64 1.0 - 8308.999999...	condition object A 13.4% 2 others 8.6% Missing 78%	structureNr float64 100545.0 - 121361.0
0	N1	nan	Others	LRPS	Start of Road afte...	814	nan	nan
1	N1	nan	Culvert	LRPSa	Box Culvert	8	nan	nan
2	N1	nan	CrossRoad	LRPSb	Intersection with ...	178	nan	nan
3	N1	nan	KmPost	LRP001	Km post missing	1000	nan	nan
0	N1	1.8	Box Culvert	LRP001a	.	11.3	A	117861
4	N1	nan	KmPost	LRP002	Km post missing	130	nan	nan
5	N1	nan	Culvert	LRP002a	Box culvert	870	nan	nan
6	N1	nan	KmPost	LRP003	Km post missing	1000	nan	nan
7	N1	nan	KmPost	LRP004	Km post missing	175	nan	nan
8	N1	nan	SideRoad,Right	LRP004a	Road to Narayan...	825	nan	nan

745 rows, showing 10 per page Page 1 of 75

```
total_df = total_df.reset_index(drop = True)
total_df
```

	road object N1 100%	km float64 39.304 - 248.499	type object RCC Girder ... 52.6% Box Culvert 26.3% 3 others 21.1%	lrp object LRP040b 5.3% LRP043a 5.3% 17 others 89.5%	name object Baldhakhal B... 5.3% SHAHID NAG... 5.3% 17 others 89.5%	length float64 1.0 - 98.87	condition object C C 100%	structureNr float64 100583.0 - 121136.0
138	N1	39.304	RCC Girder Bridge	LRP040b	Baldhakhal Bridge	23.1	C	120035
144	N1	42.936	RCC Girder Bridge	LRP043a	SHAHID NAGIR	16.5	C	112552
150	N1	45.164	Box Culvert	LRP046a	SARKAR BARI CU...	3.9	C	101134
155	N1	46.871	Slab Culvert	LRP047a	AMIRABATH SLA...	6.5	C	101140
160	N1	48.384	RCC Bridge	LRP049a	Ginlatoly	15.9	C	112549
169	N1	52.313	Box Culvert	LRP053a	AITAL BANGA BO...	4.5	C	101151
174	N1	56.006	RCC Girder Bridge	LRP056a	Illiot bazar Bridge	28.9	C	120394
207	N1	71.156	RCC Girder Bridge	LRP072a	Chandinan-Shah...	22.55	C	120395
223	N1	78.382	Slab Culvert	LRP079a	KABILA DUBARC...	1	C	101191
256	N1	93.6	RCC Girder Bridge	LRP094a	Mostapur Bridge	27.35	C	120400

19 rows, showing 10 per page Page 1 of 2

```
# The model used an id column, so we decided to create one based on the index
total_df['id'] = total_df.index
total_df
```

	road object N1 100%	km float64 1.8 - 253.777	type object KmPost 31.9% Culvert 20.5% 17 others 47.5%	lrp object LRP008b 0.3% LRP010a 0.3% 622 others 99.5%	name object Km post mi... 21.2% Box culvert 19.7% 276 others 59.1%	length float64 1.0 - 8308.999999...	condition object A 13.4% 2 others 8.6% Missing 78%	structureNr float64 100545.0 - 121361.0
0	N1	nan	Others	LRPS	Start of Road afte...	814	nan	nan
1	N1	nan	Culvert	LRPSa	Box Culvert	8	nan	nan
2	N1	nan	CrossRoad	LRPSb	Intersection with ...	178	nan	nan
3	N1	nan	KmPost	LRP001	Km post missing	1000	nan	nan
4	N1	1.8	Box Culvert	LRP001a	.	11.3	A	117861
5	N1	nan	KmPost	LRP002	Km post missing	130	nan	nan
6	N1	nan	Culvert	LRP002a	Box culvert	870	nan	nan
7	N1	nan	KmPost	LRP003	Km post missing	1000	nan	nan
8	N1	nan	KmPost	LRP004	Km post missing	175	nan	nan
9	N1	nan	SideRoad,Right	LRP004a	Road to Narayan...	825	nan	nan

745 rows, showing 10 per page Page 1 of 75

```
#We only selected the columns needed for the python model
total_df = total_df[['road', 'id', 'model_type', 'name', 'lat', 'lon', 'length', 'condition']]
total_df
```

	road object	id int64	model_type object	name object	lat float64	lon float64	length float64	condition object D

0 rows, showing 10 per page Page 0 of 0

```
#To be sure we did it correct, we decided to look at the length of the road, which is 252194meters,
#Which corresponds to 252km.
total_df.length.sum()
```

```
252194.27000000002
```