In [2]:

```
import pandas as pd
```

In [3]:

```
# df = pd.read_csv("../data/dry_etch.csv")

# As the columns cell data contains 'whitespace'. So, just trim it while fetching data.
df = pd.read_csv("../data/dry_etch.csv",
                 converters={'EquipID': str.strip,
                             'RecipeUsedOnEquip': str.strip,
                             'ActualDateIn': str.strip,
                             'LotId': str.strip})
```

In [4]:

```
# show the column headers
df.columns.to_list()
```

Out[4]:

```
['EquipID', 'RecipeUsedOnEquip', 'ActualDateIn', 'LotId']
```

In [5]:

```
# first 5 rows
df.head()
```

Out[5]:

	EquipID	RecipeUsedOnEquip	ActualDateIn	LotId
0	ASBE1	2	09-03-2021 08:39	F20280002.F1
1	ASBE1	2	09-03-2021 05:44	F20350002.F1
2	ASBE1	2	09-03-2021 04:53	F20200003.F1
3	ASBE1	#2	08-03-2021 12:41	F20370001.F1
4	ASBE1	#2	08-03-2021 10:25	F20280002.F1

In [6]:

```
# last 5 rows
df.tail()
```

Out[6]:

	EquipID	RecipeUsedOnEquip	ActualDateIn	LotId
105583				
105584				
105585				
105586				
105587				

In [7]:

```
# dropping rows where at least one element is missing
nan_val = float("NaN")

# replace "" with "NaN"
df.replace("", nan_val, inplace=True)

# drop rows with column - "EquipID" values as "NaN"
df.dropna(subset = ["EquipID"], inplace=True)
```

In [8]:

```
df.tail()
```

Out[8]:

	EquipID	RecipeUsedOnEquip	ActualDateIn	LotId
17278	RESP1B	Wf 6. Z-VIA+ASH.1MIN; WF 7. Z-VA+ASH1MIN150W; ...	24-09-2013 12:02	S13360004.F6
17279	RESP1B	Z-VIA+ASH	21-09-2013 11:51	S13360004.F5
17280	RESP1B	BARC+ME120SEC	20-09-2013 15:20	S13360004.F4
17281	RESP1B	See the Remark	20-09-2013 15:18	S13360004.F2
17282	RESP1B	SPACER-ETCH	20-04-2013 12:09	T005547H1.F1

In [9]:

```
# create new CSV file
# purpose: to view & apply filter while viewing in Excel
df.to_csv('out.csv', index=False)
```

In [10]:

```
# Fetch all the unique equip-ids provided in the data using `set()`
equip_list_unique = list(set(df["EquipID"].to_list()))
equip_list_unique
['CLMUV1',
 'REOX1B',
 'CLREPL1',
 'ASFE1',
 'CLREML1',
 'REPL1B',
 'REOX1A',
 'RESP1A',
 'ASBE1',
 'REML1C,ASBE1',
 'ASFE1,ASBE1',
 'REPL1A',
 'REML1C',
 'ASBE1,REML1C',
 'CLFRBG1',
 'CLREOX1',
 'RESP1B',
 'REOX1C',
 'CLRESP1',
 'REML1A']
equip_list_unique
```

Out[10]:

```
['ASFE1,ASBE1',
 'ASFE1',
 'CLMUV1',
 'CLREPL1',
 'REOX1C',
 'REPL1B',
 'REML1A',
 'REOX1A',
 'ASBE1,REML1C',
 'REML1C,ASBE1',
 'REPL1A',
 'CLFRBG1',
 'ASBE1',
 'REOX1B',
 'RESP1A',
 'CLREOX1',
 'CLRESP1',
 'RESP1B',
 'CLREML1',
 'REML1C']
```

Recipes run in ASBE1

Objectives

- show the recipes run in lots

- frequency of each recipe

In [11]:

```
# show the equip names by which the entry has been done in DMIS
asbe1_equip_names = [e for e in equip_list_unique if "ASBE1" in e]
asbe1_equip_names
```

Out[11]:

```
['ASFE1,ASBE1', 'ASBE1,REML1C', 'REML1C,ASBE1', 'ASBE1']
```

In [12]:

```
# create df for "EquipID" col values as `asbe1_equip_names` list
df_asbe1 = df[df["EquipID"].isin(asbe1_equip_names)]

# first 5 rows
df_asbe1.head()
```

Out[12]:

	EquipID	RecipeUsedOnEquip	ActualDateIn	LotId
0	ASBE1	2	09-03-2021 08:39	F20280002.F1
1	ASBE1	2	09-03-2021 05:44	F20350002.F1
2	ASBE1	2	09-03-2021 04:53	F20200003.F1
3	ASBE1	#2	08-03-2021 12:41	F20370001.F1
4	ASBE1	#2	08-03-2021 10:25	F20280002.F1

In [13]:

```
# fetch all the unique recipes in this equipment/chamber by `asbe1_equip_names` list
recipes_asbe1 = list(set(df_asbe1["RecipeUsedOnEquip"].to_list()))
recipes_asbe1
```

Out[13]:

```
[': Recipe # 2',
 'Recipe: Recipe # 2',
 '3',
 '#2',
 'Recipe#6',
 'Recipe # 4',
 '2',
 'TS18SLSTRIP-REML1',
 'Recipe#7(300sec) in ASFE1,Recipe:TS18SLStrip in ASP Chamber',
 'test',
 'See Remark',
 'Recipe#1',
 '#6',
 '# 2',
 'ipe # B( if in ASBE1); Recipe:TS18SLSTRIP(if in ASP Chamber of CLREML
1).',
 'Recipe # 2',
 '# 3',
```

Recipes run in ASFE1

Objectives

- show the recipes run in lots
- frequency of each recipe

In [14]:

```
# show the equip names by which the entry has been done in DMIS
asfe1_equip_names = [e for e in equip_list_unique if "ASFE1" in e]
asfe1_equip_names
```

Out[14]:

```
['ASFE1,ASBE1', 'ASFE1']
```

In [15]:

```
# create df for "EquipID" col values as `asbe1_equip_names` list
df_asfe1 = df[df["EquipID"].isin(asfe1_equip_names)]

# first 5 rows
df_asfe1.head()
```

Out[15]:

	EquipID	RecipeUsedOnEquip	ActualDateIn	LotId
1011	ASFE1	2	16-03-2021 12:07	F20370001.F1
1012	ASFE1	#2	15-03-2021 21:14	F19410001.F1
1013	ASFE1	#2	15-03-2021 19:33	F21040001.F1
1014	ASFE1	#2	15-03-2021 16:41	F20280002.F1
1015	ASFE1	D	15-03-2021 10:13	F21110001.F1

In [16]:

```
# fetch all the unique recipes in this equipment/chamber by `asbe1_equip_names` list
recipes_asfe1 = list(set(df_asfe1["RecipeUsedOnEquip"].to_list()))
recipes_asfe1
```

Out[16]:

```
['ASHER1',
 'RAA_FI"',
 ': Recipe # 2',
 'Recipe: Recipe # 2',
 'NO  Run',
 '3',
 '2,5,and 1',
 'B D',
 '#2',
 'Recipe: #2',
 'Recipe 2',
 'Recipe2',
 'Z-SIPMGC1',
 'C1D_ASHING-NFM--Post',
 'sorter.  Recipe: Recipe # D',
 'Recipe # 4',
 '2',
 'Recipe: Recipe # 3',
```

In [ ]:

```

```

Recipes run in REML1A

Recipes run in ASBE1

Recipes run in ASBE1

In [ ]:

```

```