In [11]:
```
import matplotlib as plt
import seaborn as sns
import pandas as pd
import numpy as np
import json as js
from collections import Counter
```

In [2]:
```
pwd
```

Out[2]:
```
'/Users/mingyuanma/Desktop'
```

In [6]:
```
before = pd.read_csv("before_joint_dim.csv")
before = before.loc[:,["ResponseId", "label_dim1", "label_dim2"]].groupby(
                "ResponseId").agg(lambda x: x.iloc[0])
before
```

Out[6]:

ResponseId	label_dim1	label_dim2
R_0uOsLe6BeLnUee5	3	3
R_10CNCLuGIedQoz2	1	3
R_12nnne04OpWcY9q	2	6
R_1C3aMTZKkeXGDUk	2	4
R_1DS9jhDoVpEt1m2	1	3
...	...	...
R_xA5B0rSmbSjH6XT	2	3
R_yjpyDhIt4Ei7bfX	2	4
R_yw09VUoOXdh6KA1	6	3
R_yyU2mNdEqpH2jWp	8	5
R_zTlbYwl8ijcujhn	1	7

170 rows × 2 columns

In [7]:
```
after = pd.read_csv("after_joint_dim.csv")
after = after.loc[:,["ResponseId", "label_dim1", "label_dim2"]].groupby(
                "ResponseId").agg(lambda x: x.iloc[0])
after
```

Out[7]:

ResponseId	label_dim1	label_dim2
R_0HBxBV8U696D9QJ	2	4
R_0SZ9OfuWvmqeMqR	6	5
R_10DdLv5uPLAfpAR	3	1
R_10Ggl6dEPfo9ipw	1	3
R_10HkPUkR6o0qDFT	9	1
...	...	...
R_yZJ03FFfe8jfazn	9	3
R_ykkhkGYu1KpIM8h	1	3
R_ym9gyf6T2XORxAt	1	3
R_z8OxWJ4ua6nKKB3	8	9
R_zZQIbE0LFD13yRX	3	5

245 rows × 2 columns

In [39]:
```
Counter(before[before["label_dim1"] == 1]["label_dim2"])
```

Out[39]:
```
Counter({3: 33, 4: 3, 1: 7, 5: 2, 8: 1, 7: 2})
```

In [40]:
```
def get_distribution(df):
    types_1 = []
    num = []
    types_2 = []
    for i in np.arange(1, 10):
        counting = Counter(df[df["label_dim1"] == i]["label_dim2"])
        for j in np.arange(1, 10):
            types_1.append(i)
            types_2.append(j)
            num.append(counting[j])
    d = {
        "dim1": types_1,
        "dim2": types_2,
        "frequency": num
    }
    return pd.DataFrame(d)
```

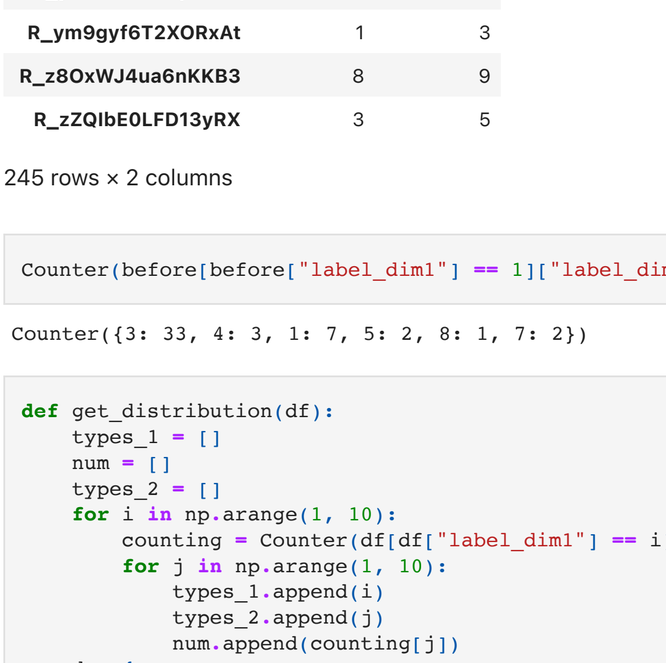
In [42]:
```
df = get_distribution(before)
df
```

Out[42]:

	dim1	dim2	frequency
0	1	1	7
1	1	2	0
2	1	3	33
3	1	4	3
4	1	5	2
...	...	...	...
76	9	5	0
77	9	6	0
78	9	7	0
79	9	8	0
80	9	9	0

81 rows × 3 columns

In [46]:
```
sns.catplot(x="dim2", y="frequency",
                hue="dim1",
                data=df, kind="bar", alpha=0.5);
```



In [ ]:

In [ ]:

In [ ]:

In [ ]: