

Finish Time Groups

In [1]:

```
import pandas as pd
import numpy as np
from scipy.stats import chi2_contingency
from scipy.stats import fisher_exact
data = pd.read_csv(filepath_or_buffer='.././../Archive/HTWTempRatios.csv')
```

15 Min time groups

In [2]:

```
#Compute a contingency table for age groups hitting the wall.
data["HTW"] = (data['DoS15km'] >= 0.25) | (data['DoS20km'] >= 0.25)
data["FTGroup"] = 0

data.loc[(data['Time'] >= 45*60) & (data['Time'] < 75*60), 'FTGroup'] = 1 #remove any with missing/unrealistic time
data.loc[(data['Time'] >= 75*60) & (data['Time'] < 90*60), 'FTGroup'] = 2
data.loc[(data['Time'] >= 90*60) & (data['Time'] < 105*60), 'FTGroup'] = 3
data.loc[(data['Time'] >= 105*60) & (data['Time'] < 120*60), 'FTGroup'] = 4
data.loc[(data['Time'] >= 120*60) & (data['Time'] < 135*60), 'FTGroup'] = 5
data.loc[(data['Time'] >= 135*60) & (data['Time'] < 150*60), 'FTGroup'] = 6
data.loc[(data['Time'] >= 150*60) & (data['Time'] < 165*60), 'FTGroup'] = 7
data.loc[(data['Time'] >= 165*60) & (data['Time'] < 180*60), 'FTGroup'] = 8
data.loc[data["Time"] >= 180*60, 'FTGroup'] = 9

ctab = pd.crosstab(data["FTGroup"], data['HTW'])
f_ctab = pd.crosstab((data.loc[data["Gender"] == "F"])["FTGroup"], data['HTW'])
m_ctab = pd.crosstab((data.loc[data["Gender"] == "M"])["FTGroup"], data['HTW'])
ctab
```

Out[2]:

	HTW	False	True
FTGroup			
1	594	1	
2	11321	37	
3	62201	831	
4	123745	4701	
5	99970	9649	
6	53992	10185	
7	22463	6052	
8	8424	3105	
9	4445	1780	

In [3]:

```
c, p, dof, expected = chi2_contingency(ctab)
print("Chi-square HTW (all runners) per 15-min finish time group p: ", p)
```

Chi-square HTW (all runners) per 15-min finish time group p: 0.0

In [4]:

```
f_ctab
```

Out[4]:

	HTW	False	True
FTGroup			
1	62	0	
2	519	0	
3	7039	24	
4	32408	217	
5	43086	1078	
6	29039	2155	
7	13169	1977	
8	5079	1263	
9	2473	821	

In [5]:

```
m_ctab
```

Out[5]:

	HTW	False	True
FTGroup			
1	532	1	
2	10802	37	
3	55162	807	
4	91337	4484	
5	56884	8571	
6	24953	8030	
7	9294	4075	
8	3345	1842	
9	1972	959	

In [6]:

```
c, p, dof, expected = chi2_contingency(f_ctab)
print("Chi-square HTW (female) per 15-min finish time group p: ", p)
c, p, dof, expected = chi2_contingency(m_ctab)
print("Chi-square HTW (male) per 15-min finish time group p: ", p)
```

Chi-square HTW (female) per 15-min finish time group p: 0.0  
Chi-square HTW (male) per 15-min finish time group p: 0.0

In [7]:

```
f_ctab2 = pd.crosstab((data.loc[data["Gender"] == "F"])["FTGroup"], (data['SplitRatio'] <= 1))
m_ctab2 = pd.crosstab((data.loc[data["Gender"] == "M"])["FTGroup"], (data['SplitRatio'] <= 1))
ctab2 = pd.crosstab(data["FTGroup"], (data['SplitRatio'] <= 1))
ctab2
```

Out[7]:

	SplitRatio	False	True
FTGroup			
1	565	30	
2	9668	1690	
3	53137	9895	
4	112058	16388	
5	100379	9240	
6	61133	3044	
7	27797	718	
8	11256	273	
9	6017	208	

In [8]:

```
c, p, dof, expected = chi2_contingency(ctab2)
print("Chi-square (neg split) for all runners per 15-min finish time group p: ", p)
```

Chi-square (neg split) for all runners per 15-min finish time group p: 0.0

In [9]:

```
f_ctab2
```

Out[9]:

	SplitRatio	False	True
FTGroup			
1	62	0	
2	484	35	
3	6054	1009	
4	27905	4720	
5	39549	4615	
6	29304	1890	
7	14713	433	
8	6187	155	
9	3180	114	

In [10]:

```
m_ctab2
```

Out[10]:

	SplitRatio	False	True
FTGroup			
1	503	30	
2	9184	1655	
3	47083	8886	
4	84153	11668	
5	60830	4625	
6	31829	1154	
7	13084	285	
8	5069	118	
9	2837	94	

In [11]:

```
c, p, dof, expected = chi2_contingency(f_ctab2)
print("Female (neg split) per 15-min finish time group p: ", p)

c, p, dof, expected = chi2_contingency(m_ctab2)
print("Male (neg split) per 15-min finish time group p: ", p)
```

Female (neg split) per 15-min finish time group p: 0.0  
Male (neg split) per 15-min finish time group p: 0.0

In [ ]: