

Friedman_hw2

October 7, 2018

```
In [2]: from __future__ import print_function
```

```
import scipy.stats as st
import numpy as np
import pandas as pd
```

0.1 Problem 5

```
In [4]: # Load baseball data from .csv
df = pd.read_csv('baseball.csv')
df
```

```
Out[4]:
```

	FirstName	LastName	At-Bats	Hits	BattingAverage	RemainingAt-Bats	\
0	Roberto	Clemente	45	18	0.400	367	
1	Frank	Robinson	45	17	0.378	426	
2	Frank	Howard	45	16	0.356	521	
3	Jay	Johnstone	45	15	0.333	275	
4	Ken	Berry	45	14	0.311	418	
5	Jim	Spencer	45	14	0.311	466	
6	Don	Kessinger	45	13	0.289	586	
7	Luis	Alvarado	45	12	0.267	138	
8	Ron	Santo	45	11	0.244	510	
9	Ron	Swaboda	45	11	0.244	200	
10	Rico	Petrocelli	45	10	0.222	538	
11	Ellie	Rodriguez	45	10	0.222	186	
12	George	Scott	45	10	0.222	435	
13	Del	Unser	45	10	0.222	277	
14	Billy	Williams	45	10	0.222	591	
15	Bert	Campaneris	45	9	0.200	558	
16	Thurman	Munson	45	8	0.178	408	
17	Max	Alvis	45	7	0.156	70	

	RemainingAverage	SeasonAt-Bats	SeasonHits	SeasonAverage
0	0.3460	412	145	0.352
1	0.2981	471	144	0.306
2	0.2764	566	160	0.283
3	0.2218	320	76	0.238
4	0.2727	463	128	0.276