SC_hw1

R26131133

titanic
12 Variables   891 Observations

PassengerId

n	missing	distinct	Info	Mean	pMedian	Gmd	.05	.10	.25	.50	.75	.90	.95
891	0	891	1	446	446	297.3	45.5	90.0	223.5	446.0	668.5	802.0	846.5

```
lowest :  1   2   3   4   5, highest: 887 888 889 890 891
```

Survived

n	missing	distinct	Info	Sum	Mean
891	0	2	0.71	342	0.3838

Pclass

n	missing	distinct	Info	Mean	pMedian	Gmd
891	0	3	0.81	2.309	2.5	0.8631

```
Value           1     2     3
Frequency     216   184   491
Proportion  0.242 0.207 0.551
```

For the frequency table, variable is rounded to the nearest 0

Name

n	missing	distinct
891	0	891

```
lowest : Abbing, Mr. Anthony              Abbott, Mr. Rossmore Edward         Abbott, Mrs. Stanton (Rosa Hunt)    Abelson, Mr
highest: Yousseff, Mr. Gerious            Yrois, Miss. Henriette ("Mrs Harbeck") Zabour, Miss. Hileni                Zabour, Miss
```

Sex

n	missing	distinct
891	0	2

```
Value      female    male
Frequency     314     577
Proportion  0.352   0.648
```

Age

n	missing	distinct	Info	Mean	pMedian	Gmd	.05	.10	.25	.50	.75	.90	.95
714	177	88	0.999	29.7	29	16.21	4.00	14.00	20.12	28.00	38.00	50.00	56.00

```
lowest : 0.42 0.67 0.75 0.83 0.92, highest: 70   70.5 71   74   80
```

SibSp

n	missing	distinct	Info	Mean	pMedian	Gmd
891	0	7	0.669	0.523	0.5	0.823

Value	0	1	2	3	4	5	8
Frequency	608	209	28	16	18	5	7
Proportion	0.682	0.235	0.031	0.018	0.020	0.006	0.008

For the frequency table, variable is rounded to the nearest 0

Parch

n	missing	distinct	Info	Mean	pMedian	Gmd
891	0	7	0.556	0.3816	0	0.6259

Value	0	1	2	3	4	5	6
Frequency	678	118	80	5	4	5	1
Proportion	0.761	0.132	0.090	0.006	0.004	0.006	0.001

For the frequency table, variable is rounded to the nearest 0

Ticket

n	missing	distinct
891	0	681

```
lowest : 110152    110413    110465    110564    110813
highest: W./C. 6608  W./C. 6609  W.E.P. 5734  W/C 14208   WE/P 5735
```

Fare

n	missing	distinct	Info	Mean	pMedian	Gmd	.05	.10	.25
891	0	248	1	32.2	19.6	36.78	7.225	7.550	7.910

.50	.75	.90	.95
14.454	31.000	77.958	112.079

```
lowest : 0       4.0125  5       6.2375  6.4375 , highest: 227.525 247.521 262.375 263    512.329
```

Cabin

n	missing	distinct
204	687	147

```
lowest : A10 A14 A16 A19 A20, highest: F33 F38 F4  G6  T
```

Embarked

n	missing	distinct
889	2	3

Value	C	Q	S
Frequency	168	77	644
Proportion	0.189	0.087	0.724

n: number of the data that has be observed
missing: number of missing value(NA)
distinct: number of distinct value
Info: Information Value(IV)
pMedian: pseudo-median
Gmd: Gini Mean Difference