
H2O.ai

Numeric Data Summaries

- h2o_frame[x].cor(y=None, na_rm=False, use=None)
- h2o_frame[x].kurtosis(na_rm=False)
- h2o_frame[x].max()
- h2o_frame[x].mean(skipna=False)
- h2o_frame[x].median(na_rm=False)
- h2o_frame[x].min()
- h2o_frame[x].prod(na_rm=False)
- h2o_frame[x].quantile(…)
- h2o_frame[x].sd(na_rm=False)
- h2o_frame[x].skewness(na_rm=False)
- h2o_frame[x].sum(skipna=True)
- h2o_frame[x].var(y=None, na_rm=False, use=None)

Group By Aggregation

```
1 grouped_data = census_data[["occupation", "education-num"]].group_by(["occupation"])
2 stats = grouped_data.count(na = "ignore").median(na = "ignore").mean(na = "ignore").sd(na = "ignore")
3 stats.get_frame()
```

occupation	nrow	median_education-num	mean_education-num	sdev_education-num
	0	9	9.25339	2.60279
Adm-clerical	3770	10	10.1135	1.69805
Armed-Forces	9	9	10.1111	2.02759
Craft-repair	4099	9	9.11076	2.03865
Exec-managerial	4066	12	11.4491	2.14321
Farming-fishing	994	9	8.60865	2.75607
Handlers-cleaners	1370	9	8.51022	2.20338
Machine-op-inspct	2002	9	8.48751	2.28528
Other-service	3295	9	8.77967	2.29966

Numeric Data Summaries

- `h2o_frame[x].cor(y=None, na_rm=False, use=None)`
- `h2o_frame[x].kurtosis(na_rm=False)`
- `h2o_frame[x].max()`
- `h2o_frame[x].mean(skipna=False)`
- `h2o_frame[x].median(na_rm=False)`
- `h2o_frame[x].min()`
- `h2o_frame[x].prod(na_rm=False)`
- `h2o_frame[x].quantile(…)`
- `h2o_frame[x].sd(na_rm=False)`
- `h2o_frame[x].skewness(na_rm=False)`
- `h2o_frame[x].sum(skipna=True)`
- `h2o_frame[x].var(y=None, na_rm=False, use=None)`