



Data Selection Example

```
1 tech_support_income = census_data[census_data["occupation"] == "Tech-support", "income"]
2 print(type(tech_support_income))
```

```
<class 'h2o.frame.H2OFrame'>
```

```
1 print(tech_support_income.table())
```

income	Count
<=50K	645
>50K	283

Missing Data Handling

- Missing Data Code
 - None
- Missing Data Filtering
 - `h2o_frame.filter_na_cols(frac=0.2)`
 - `h2o_frame.na_omit()`
- Missing Data Replacement
 - `h2o_frame.impute(...)`
 - `H2OGeneralizedLowRankEstimator`
- Missing Data Inclusion
 - `h2o_frame.insert_missing_values(fraction=0.1, seed=None)`
- Missing Data Summaries
 - `h2o_frame.nacnt()`

Data Selection Example

```
1 tech_support_income = census_data[census_data["occupation"] == "Tech-support", "income"]  
2 print(type(tech_support_income))
```

```
<class 'h2o.frame.H2OFrame'>
```

```
1 print(tech_support_income.table())
```

income	Count
<=50K	645
>50K	283