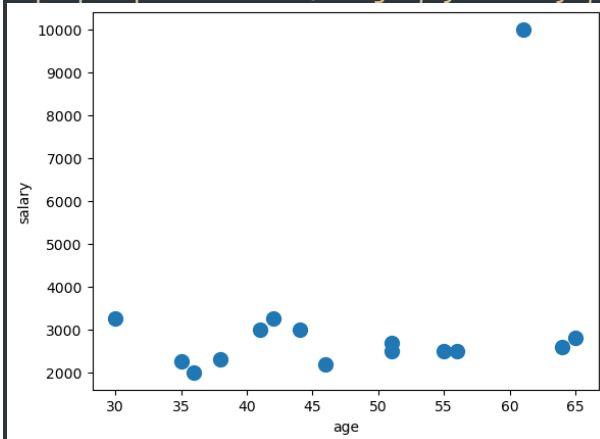


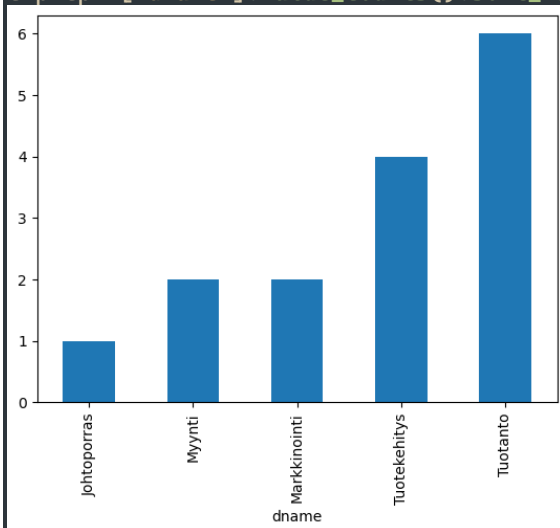
Viikko 37 -tehtävät

Tehtävä 1

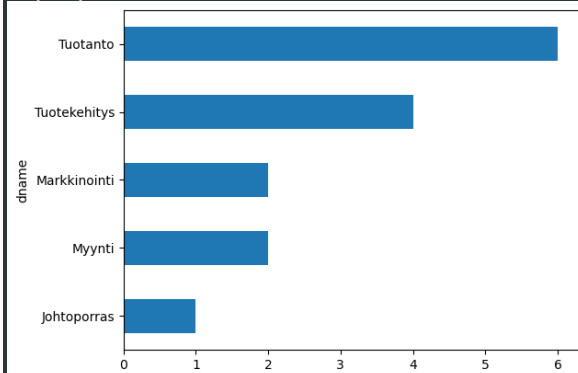
```
empDepDF = pd.read_csv('./work/viikko3/datasets/emp-dep.csv')  
empDepDF.plot.scatter(x='age', y='salary', s=100)
```



```
empDepDF['dname'].value_counts().sort_values().plot.bar()
```

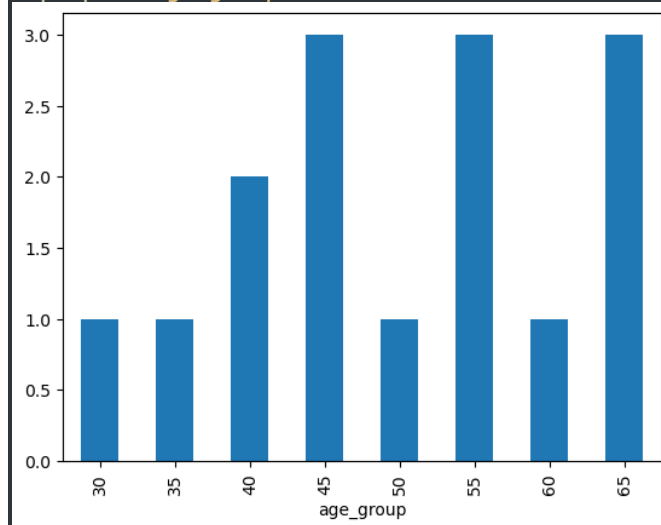


```
empDepDF['dname'].value_counts().sort_values().plot.barh()
```



Tehtävä 2

```
empDepDF['age_group'].value_counts().sort_index().plot.bar()
```



Tehtävä 3

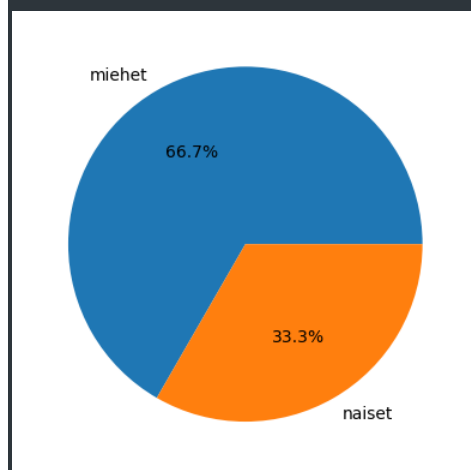
```
gender_counts = empDepDF['gender'].value_counts()
```

```
gender_mapping = {0: 'miehet', 1: 'naiset'}
```

```
gender_percentages = (gender_counts / gender_counts.sum()) * 100
```

```
gender_percentages.index = gender_percentages.index.map(gender_mapping)
```

```
plt.pie(gender_percentages, labels=gender_percentages.index, autopct='%1.1f%%')
```

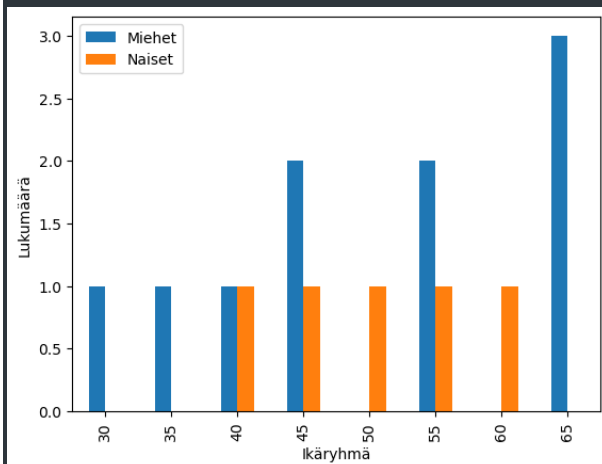


```
male_age_counts = empDepDF[empDepDF['gender'] == 0]['age_group'].value_counts()
```

```
female_age_counts = empDepDF[empDepDF['gender'] == 1]['age_group'].value_counts()
```

```
gender_counts_df = pd.DataFrame({'Miehet': male_age_counts, 'Naiset':  
female_age_counts}).fillna(0)
```

```
gender_counts_df.plot.bar(xlabel='Ikäryhmä', ylabel='Lukumäärä')
```



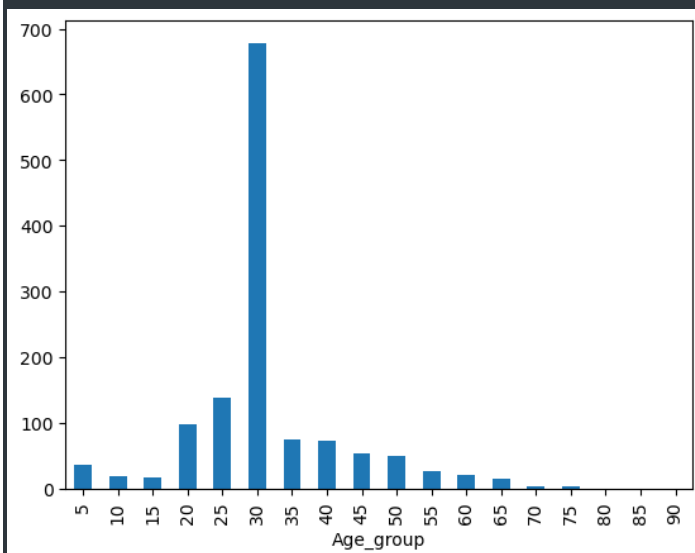
Tehtävä 4

```
titanicDF = pd.read_csv('./work/viikko3/datasets/titanic.csv')
```

```
age_bins = range(0, 95, 5)
```

```
titanicDF['Age_group'] = pd.cut(titanicDF['Age'], bins=age_bins, labels=age_bins[1:])
```

```
titanicDF['Age_group'].value_counts().sort_index().plot.bar()
```



```
male_survived_counts = titanicDF[titanicDF['Gender'] ==  
'male']['Survived'].value_counts()
```

```
female_survived_counts = titanicDF[titanicDF['Gender'] ==  
'female']['Survived'].value_counts()
```

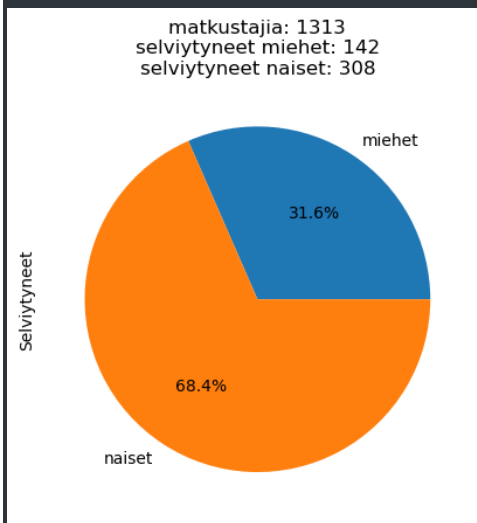
```

gender_survived_counts_df = pd.DataFrame({'miehet': male_survived_counts, 'naiset':
female_survived_counts})

gender_survived_percentages = (gender_survived_counts_df.iloc[1] /
gender_survived_counts_df.iloc[1].sum()) * 100

gender_survived_percentages.plot.pie(ylabel='Selviytyneet',
                                     autopct='%1.1f%%',
                                     title=f'matkustajia:
{titanicDF.shape[0]}\nselviytyneet miehet: {gender_survived_counts_df.iloc[1,
0]}\nselviytyneet naiset: {gender_survived_counts_df.iloc[1, 1]}'
)

```



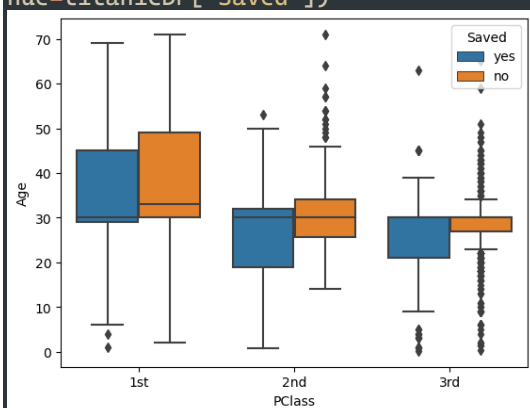
```

titanicDF = titanicDF[titanicDF['PClass'] != '*']

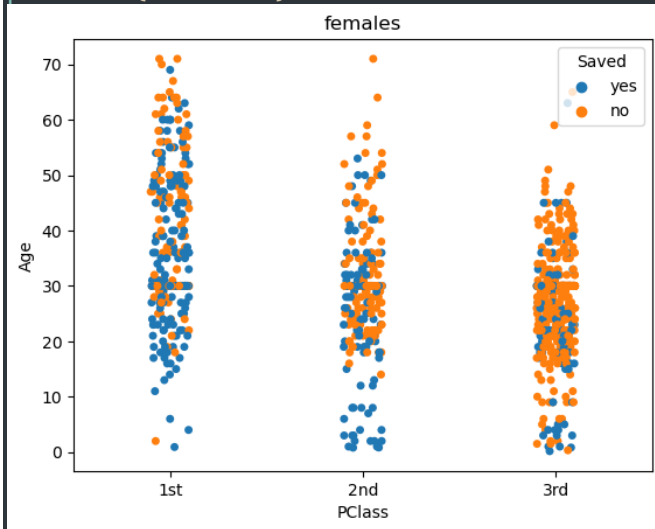
saved_mapping = {0: 'no', 1: 'yes'}
titanicDF['Saved'] = titanicDF['Survived'].map(saved_mapping)

sb.boxplot(data=titanicDF['Saved'], x=titanicDF['PClass'], y=titanicDF['Age'],
hue=titanicDF['Saved'])

```



```
sb.stripplot(data=titanicDF[titanicDF['Gender'] ==
'female']['Saved'], x=titanicDF['PClass'], y=titanicDF['Age'],
hue=titanicDF['Saved'])
plt.title("females")
```



```
sb.stripplot(data=titanicDF[titanicDF['Gender'] ==
'male']['Saved'], x=titanicDF['PClass'], y=titanicDF['Age'], hue=titanicDF['Saved'])
plt.title("males")
```

