Homework 2 - Week 9

Kelompok 7 - Citizen Data Scientist

Link Google Colab

https://colab.research.google.com/drive/1OyhW94YrXT60XYhJQM5zUujVAMfWFrd9?usp=sharing#scrollTo=wquuQRUqSkpV

Histogram

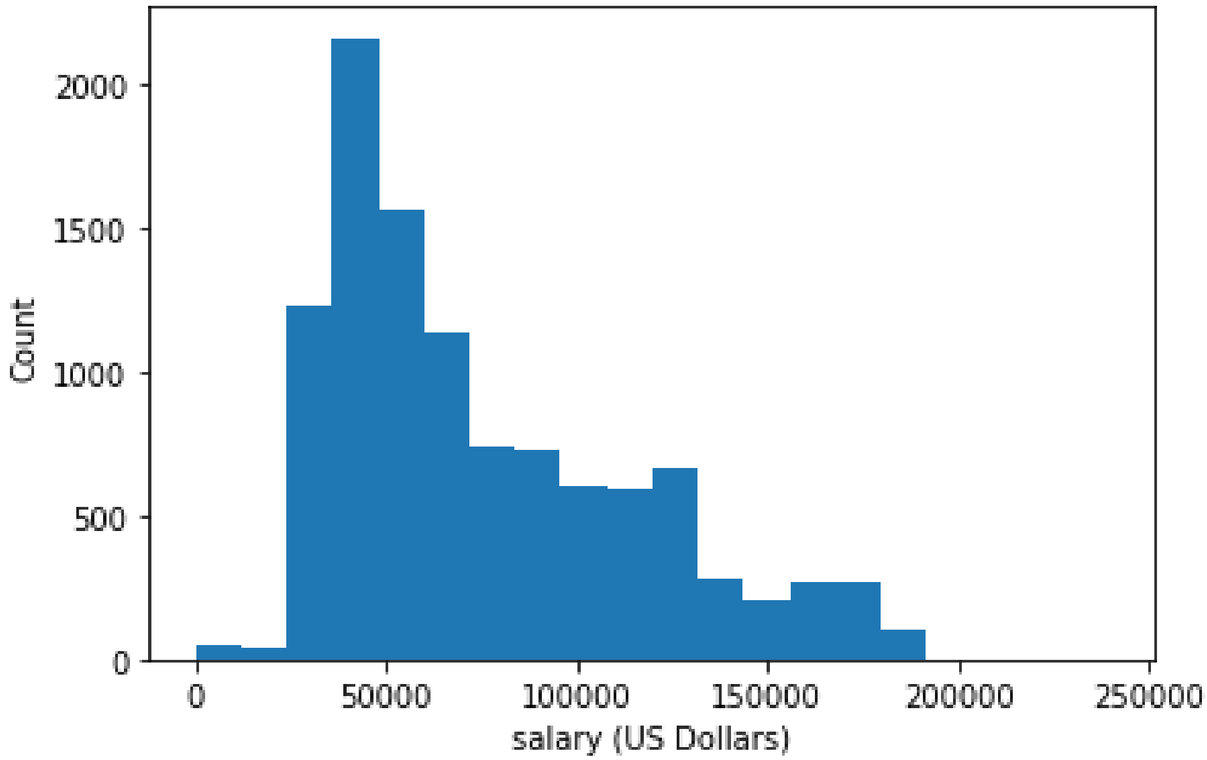
Syntax:

```
gaji = data['salary']
fig, ax = plt.subplots()

ax.hist(x=gaji, bins=20)
ax.set_xlabel('salary (US Dollars)')
ax.set_ylabel('Count')

plt.show()
```

Histogram

Result Display

```
ax.set_ylabel('Count')
```




```
ax.set_xlabel('salary (US Dollars)')
```

Bar Plot Vertikal

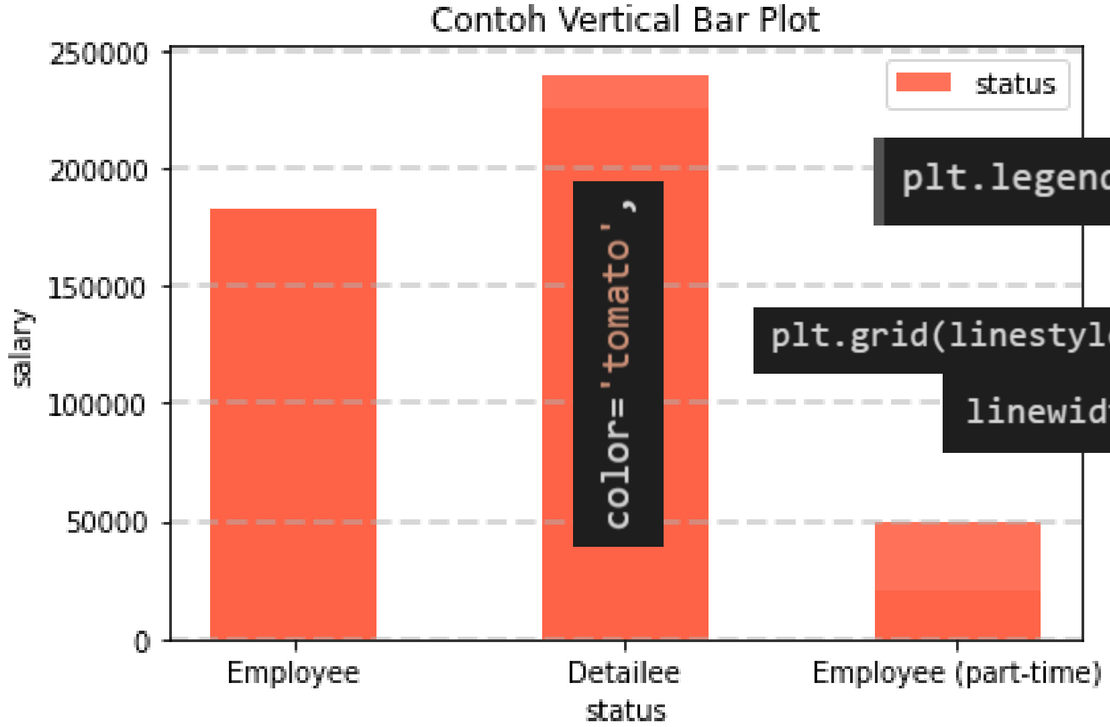
Syntax:

```
# Pengaturan Plot
# gunakan plt.bar untuk horisontal barplot
plt.bar(data.status,      # data sb. x
        data.salary,      # data sb. y
        width= 0.5,       # lebar tiap bins/persegi
        align='center',   # pengaturan rata kiri, kanan, tengah
        color='tomato',     # warna bins/persegi
        alpha=0.9)        # pengaturan transparansi

# Pengaturan Grid
plt.grid(linestyle='--',  # tipe garis grid
         linewidth=2,     # lebar garis grid
         axis='y',        # sumbu apa yang mau ditampilkan (x, y, both)
         alpha=0.5)       # transparans grid

# Pengaturan nama sumbu, nama judul plot, dana legend
plt.xlabel('status')            # nama sumbu x
plt.ylabel('salary')            # nama sumbu y
plt.title('Contoh Vertical Bar Plot')    # nama judul plot
plt.legend(labels = ['status']) # nama label

plt.show()   # memeunculkan plota
```

Bar Plot Vertikal

Result Display

```
plt.title('Contoh Vertical Bar Plot')
```




```
plt.ylabel('salary')
```

```
color='tomato',
```

```
plt.legend(labels = ['status'])
```

```
plt.grid(linestyle='--',
         linewidth=2,
```

```
plt.xlabel('status')
```

Bar Plot Horizontal

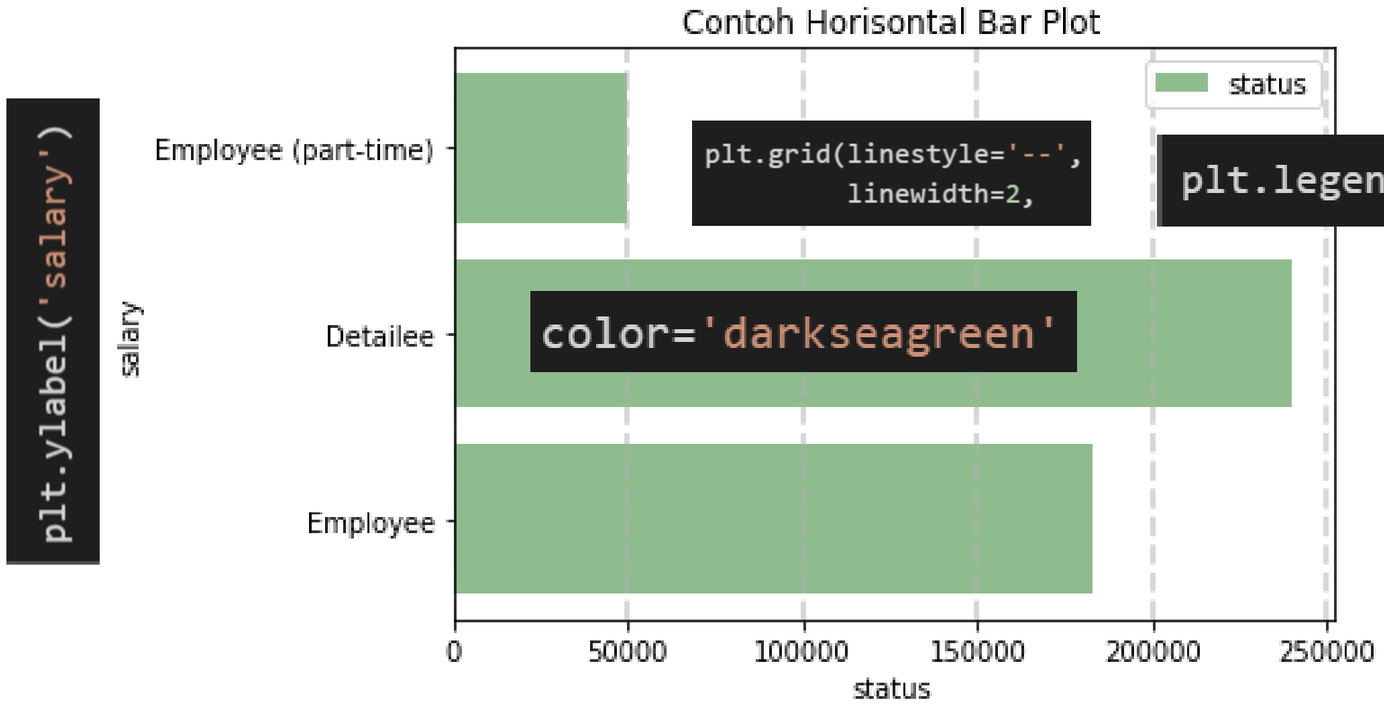
Syntax:

```
# Pengaturan Plot
# gunakan plt.barh untuk horizontal barplot
plt.barh(data.status,      # data sb. x
         data.salary,      # data sb. y
         color='darkseagreen'  # warna bar-nya
         )

# Pengaturan Grid
plt.grid(linestyle='--',  # tipe garis grid
         linewidth=2,     # lebar garis grid
         axis='x',        # sumbu apa yang mau ditampilkan (x, y, both)
         alpha=0.5)       # transparans grid

# Pengaturan nama sumbu, nama judul plot, dana legend
plt.xlabel('status')            # nama sumbu x
plt.ylabel('salary')            # nama sumbu y
plt.title('Contoh Horisontal Bar Plot')    # nama judul plot
plt.legend(labels = ['status']) # nama label

plt.show()   # memeunculkan plota
```

Bar Plot Horizontal

Result Display

```
plt.title('Contoh Horisontal Bar Plot')
```




```
plt.grid(linestyle='--',
         linewidth=2,
```

```
plt.legend(labels = ['status'])
```

```
plt.ylabel('salary')
```

```
color='darkseagreen'
```

```
plt.xlabel('status')
```

Pie Chart

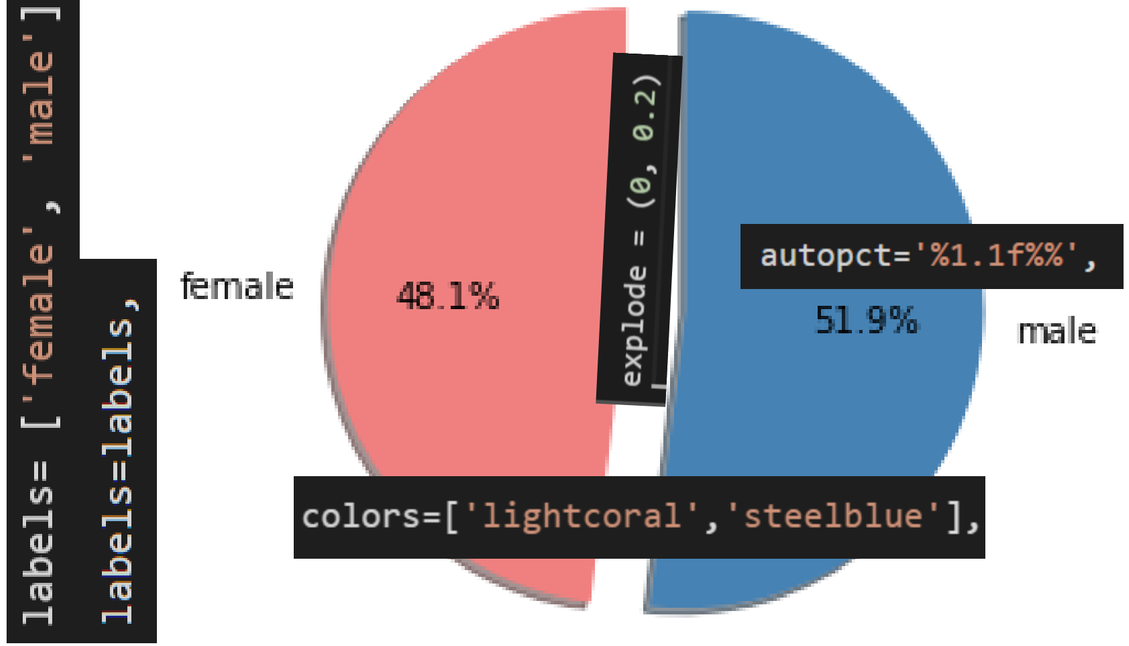
Syntax

```
data2 = data.groupby(data["gender"])["salary"].sum()    # agregasi terlebih dahulu datanya
labels= ['female', 'male']                              # tentukan label
explode = (0, 0.2)                                      # tentukan bagian mana yg mau dipisah

#mengatur piechart
plt.pie(data2,                                  # variabel yg sudah diagregasi
        colors=['lightcoral','steelblue'],      # warna
        labels=labels,                          # label
        autopct='%1.1f%%',                      # angka dalam bentuk persentase
        shadow=True,                            # memberi shadow
        explode=explode,                        # bagian untuk memisahkan
        startangle=90)                          # mulai piechart dari sudut berapa derajat

plt.show()
```

Pie Chart

Result Display

```
labels= ['female', 'male']
labels=labels,
```



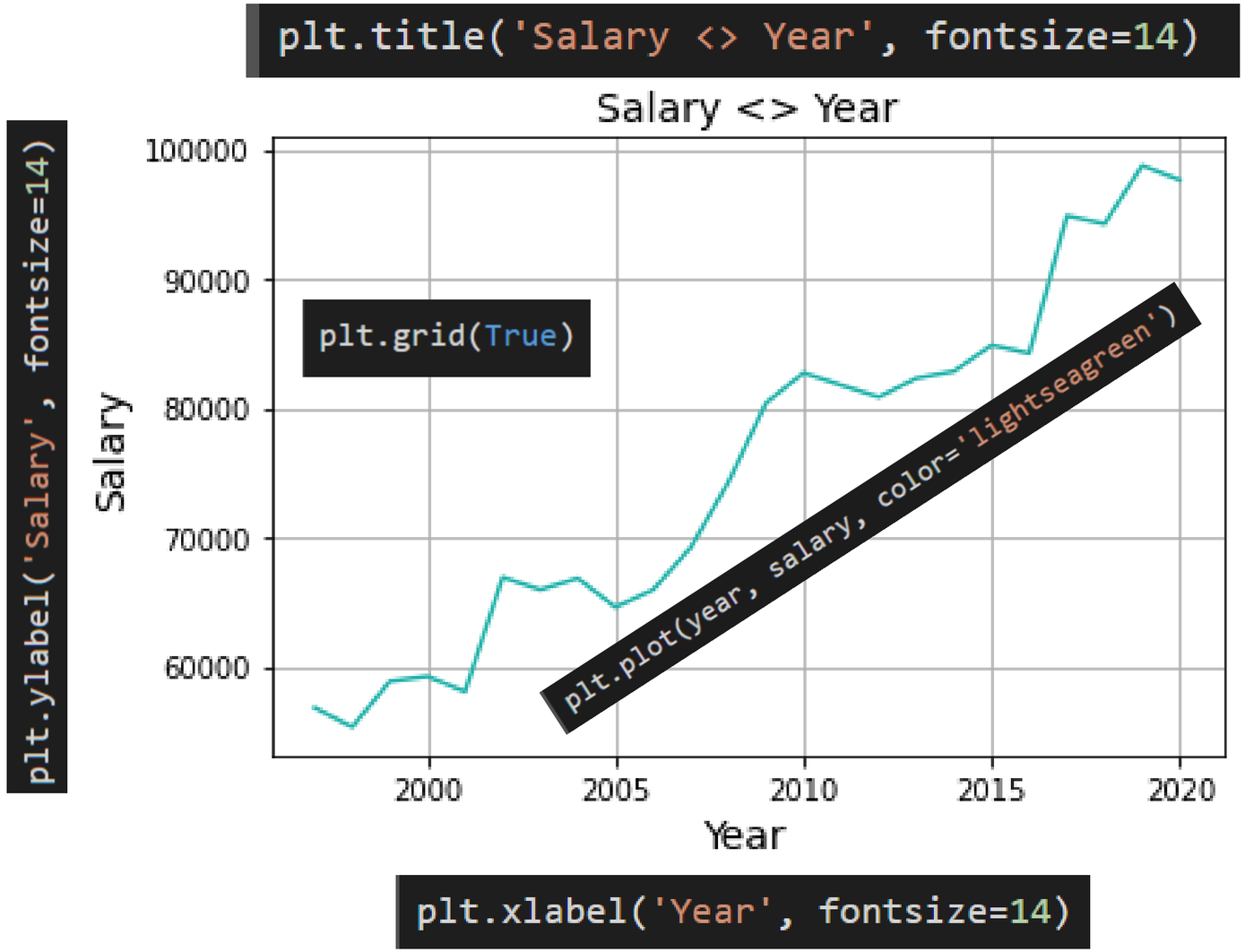
Line Chart

Syntax

```
grouped_year_salary=data.groupby(['year'])['salary'].mean().reset_index()
grouped_year_salary.columns = ['year', 'salary']

year = grouped_year_salary['year']
salary = grouped_year_salary['salary']

plt.plot(year, salary, color='lightseagreen')
plt.title('Salary <> Year', fontsize=14)
plt.xlabel('Year', fontsize=14)
plt.ylabel('Salary', fontsize=14)
plt.grid(True)
plt.show()
```

Line Chart

Result Display



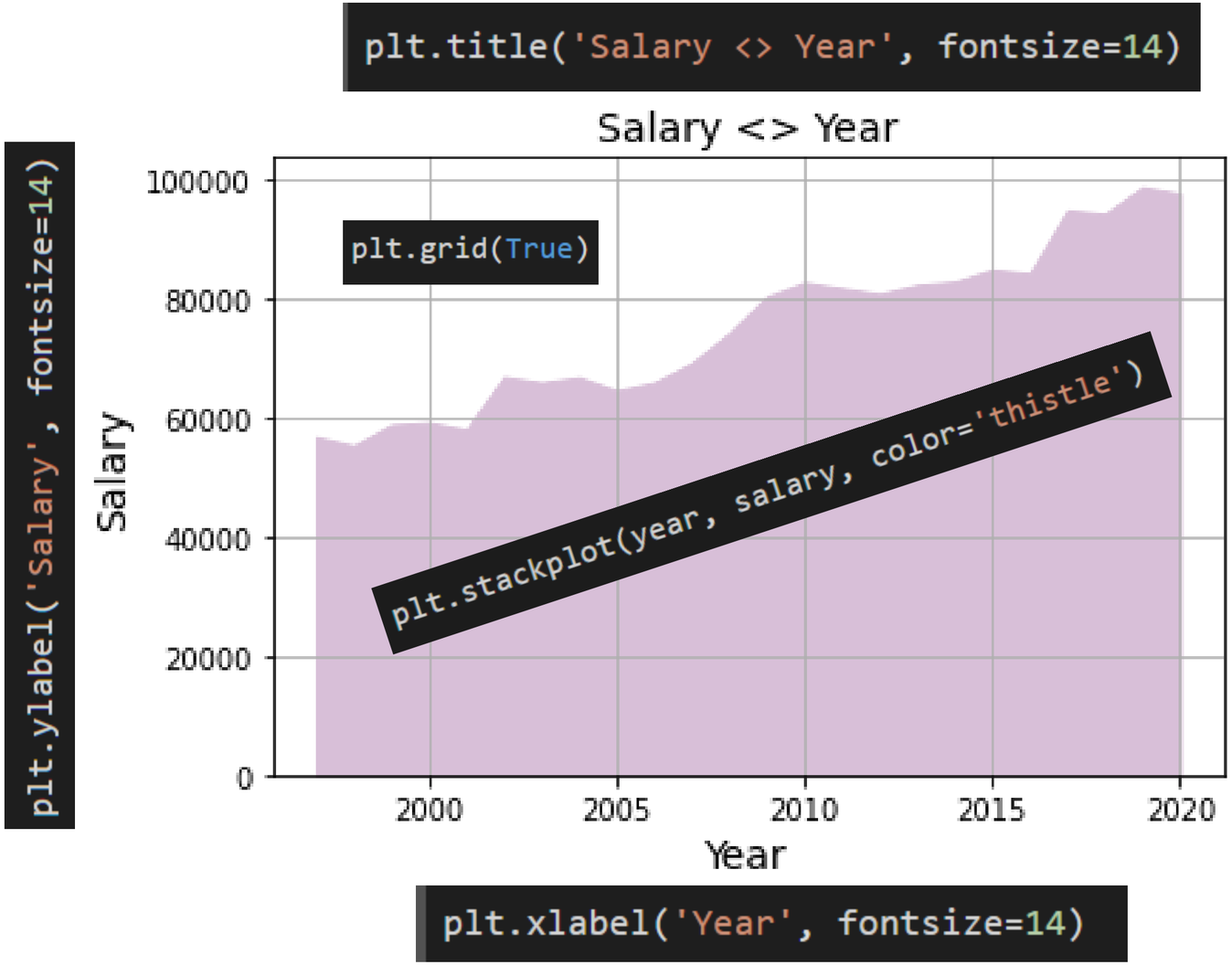
Area Plot

Syntax

```
grouped_year_salary=data.groupby(['year'])['salary'].mean().reset_index()
grouped_year_salary.columns = ['year', 'salary']

year = grouped_year_salary['year']
salary = grouped_year_salary['salary']

plt.stackplot(year, salary, color='thistle')
plt.title('Salary <> Year', fontsize=14)
plt.xlabel('Year', fontsize=14)
plt.ylabel('Salary', fontsize=14)
plt.grid(True)
plt.show()
```

Area Plot

Result Display



Scatter Plot

Syntax

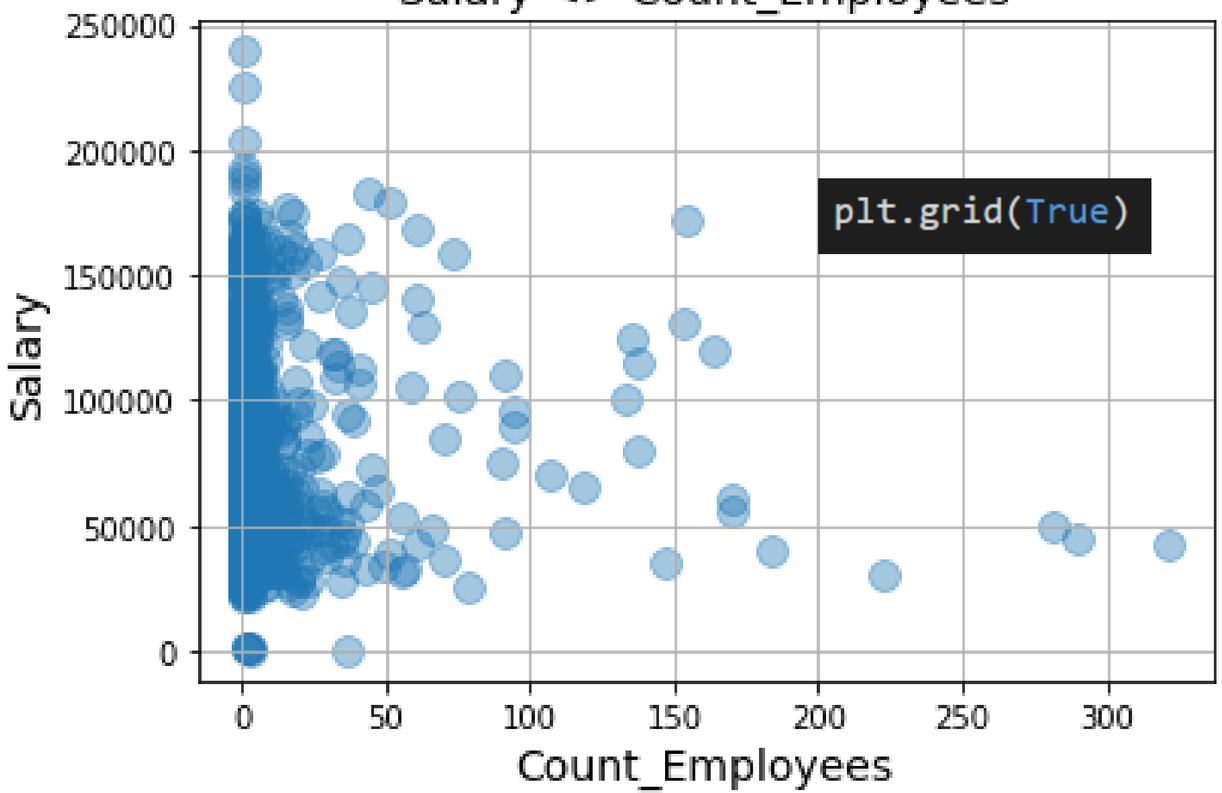
```
grouped_salary_count=data.groupby(['salary'])['name'].count().reset_index()
grouped_salary_count.columns = ['salary', 'count']
grouped_salary_count

count = grouped_salary_count['count']
salary = grouped_salary_count['salary']

plt.scatter(count, salary,alpha=0.4, s = 100)
plt.title('Salary <> Count_Employees', fontsize=14)
plt.xlabel('Count_Employees', fontsize=14)
plt.ylabel('Salary', fontsize=14)
plt.grid(True)
plt.show()
```

Scatter Plot

Result Display

```
plt.title('Salary <> Count_Employees', fontsize=14)
```

```
plt.ylabel('Salary', fontsize=14)
```




```
plt.grid(True)
```

```
plt.xlabel('Count_Employees', fontsize=14)
```