```
In [1]:  import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import seaborn as sns
         %matplotlib inline

         ger_childrens_book_authors = pd.read_csv(r"D:/cravi/Documents/Uni/DH/Statistik/Data/KinderautorenD2
         0Jhdt.tsv", sep="\t")
         ger_childrens_book_authors["Lebensspanne"] = ger_childrens_book_authors["Todesjahr"] - ger_children
         s_book_authors["Geburtsjahr"]
         ger_childrens_book_authors.head(3)
```

Out[1]:

	Nachname	Vorname	Geburtsjahr	Todesjahr	Lebensspanne
0	Abraham	Peter	1936	2015	79
1	Aichner	Fridolin	1912	1987	75
2	Alexander-Burgh	Eberhard	1929	2004	75

Barplot Lebensspanne

..der Autoren aus dem eigenen Datensatz

```
In [2]: ger_childrens_book_authors.sort_values(by = "Lebensspanne").plot.bar(figsize=(20,5), x = "Nachname"
        , y = "Lebensspanne", color ="blue", title = "Lebensspannen der Autoren")
```
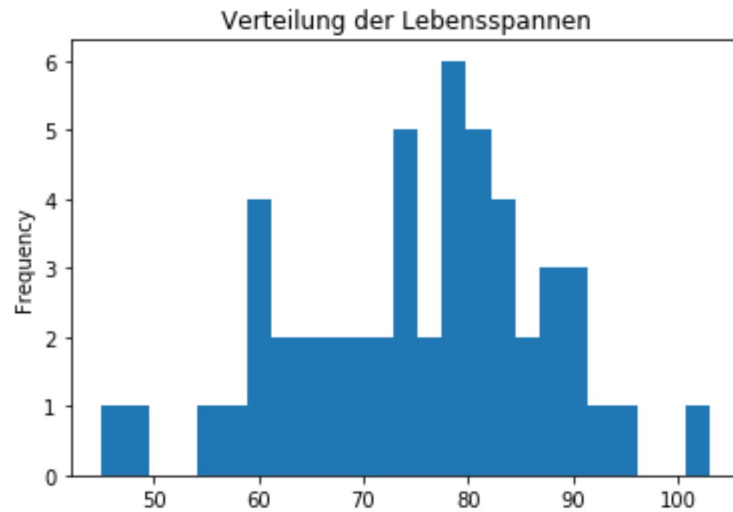
Out[2]: <matplotlib.axes._subplots.AxesSubplot at 0x200a6ec5128>
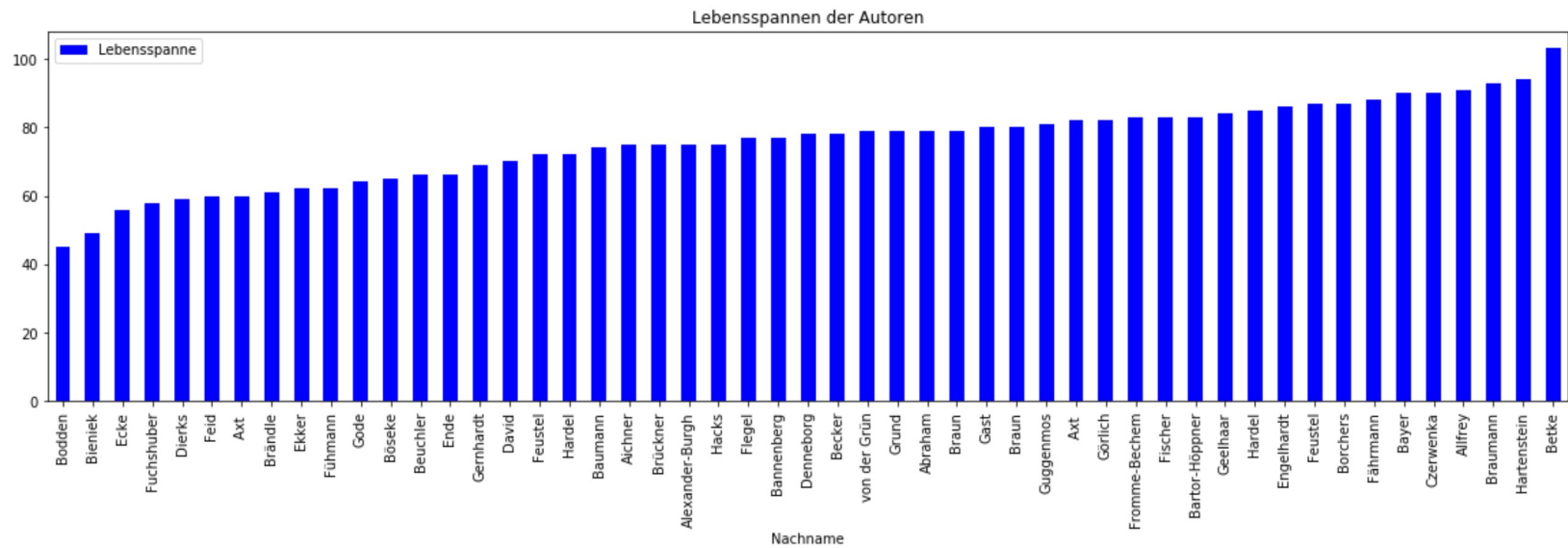
Dazugehöriges Histogramm

```
In [3]: ger_childrens_book_authors["Lebensspanne"].plot.hist(x = "Lebensspanne in Jahre", bins = 25, title
        = "Verteilung der Lebensspannen")
```
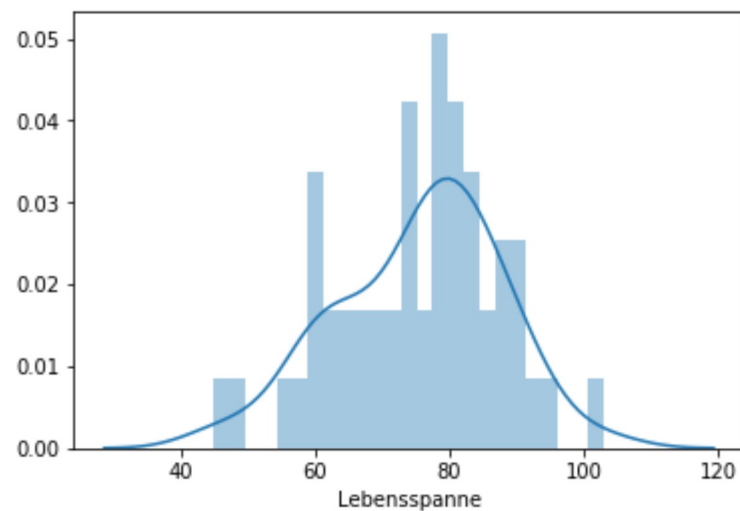
Out[3]: <matplotlib.axes._subplots.AxesSubplot at 0x200a9246080>


```
In [4]: sns.distplot(ger_childrens_book_authors['Lebensspanne'], bins=25, kde=True)
```

```
        d:\languages\python\python3.6.3\lib\site-packages\matplotlib\axes\_axes.py:6462: UserWarning: The
        'normed' kwarg is deprecated, and has been replaced by the 'density' kwarg.
          warnings.warn("The 'normed' kwarg is deprecated, and has been "
```

Out[4]: <matplotlib.axes._subplots.AxesSubplot at 0x200a9051668>

Die Lebensspannen sind annähernd Normalverteilt, die Daten sind kontinuerlich.

```
In [5]: ayl_speaker_data = pd.read_csv(r"D:/cravi/Documents/Uni/DH/Statistik/Data/AYLSpeakerData.csv", sep=
        ";")
        ayl_speaker_data.head(3)
```

Out[5]:

|	ID	label	gender	per_ms_sps	role	importance
0	0	orlando	male	120	protagonist	primary
1	1	oliver	male	40	antagonist	primary
2	2	second brother	male	2	other	minor

Barplot der per_ms_sps

```
In [6]: ayl_speaker_data.sort_values(by = "per_ms_sps", ascending = False).plot.bar(figsize=(20,5), x = "la
        bel", y = "per_ms_sps", color = "blue", title = "Redebeiträge pro Charakter")
        plt.xlabel("Character Label")
        plt.ylabel("Occurence of Speech")
```

Out[6]: Text(0,0.5,'Occurence of Speech')
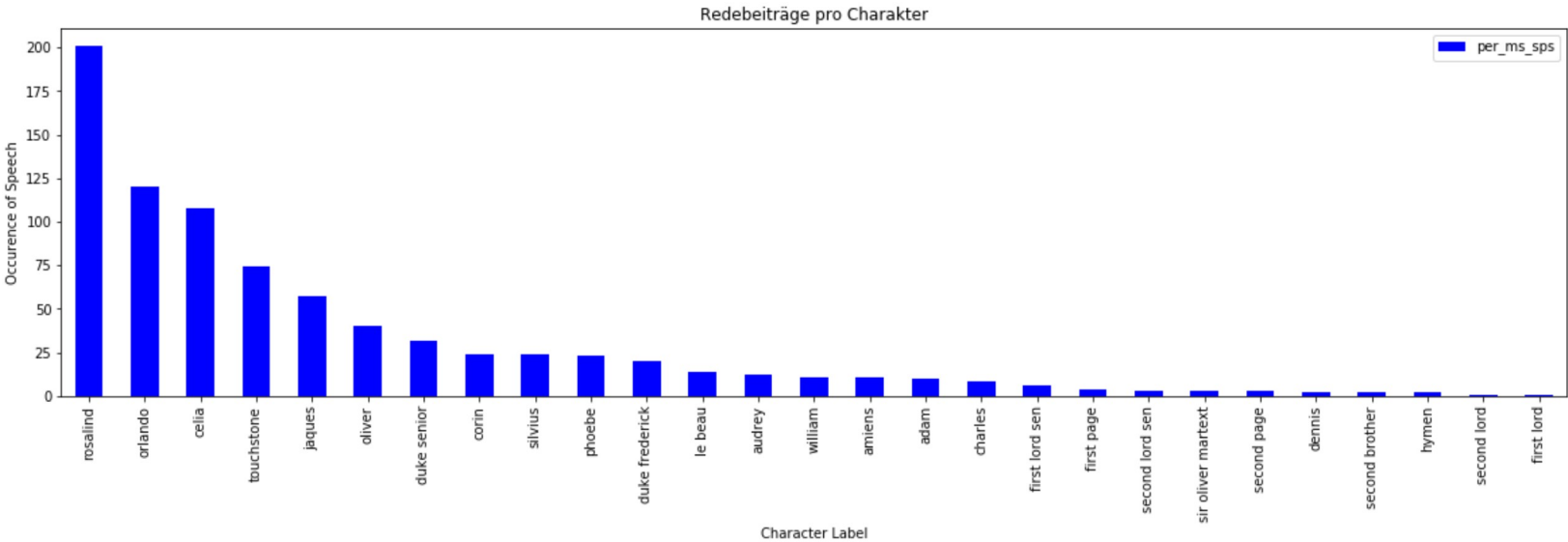
```
In [7]: ayl_speaker_data["per_ms_sps"].plot.box()
```

Out[7]: <matplotlib.axes._subplots.AxesSubplot at 0x200a925fe48>
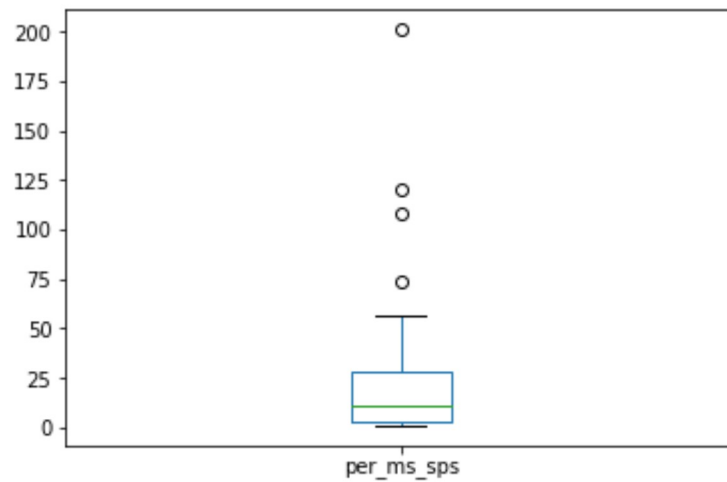
Der Barplot folgt der Zipf-Verteilung. Die gezählten Ereignisse sind nicht unabhängig voneinander, deswegen liegt keine Poisson-Verteilung vor.