

# Motorcycle Performance & Design Analysis

**Research Question:** Welche Faktoren beeinflussen die Leistung und das Design von Motorrädern?

Datensatz: Kaggle – Motorcycle Technical Specifications (1970–2022)

## 1. Imports

```
In [26]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

## 2. Daten laden

```
In [27]: df = pd.read_csv("data/all_bikez_curated.csv", low_memory=False)
df.sample(n=5)
```

Out[27]:

	Brand	Model	Year	Category	Rating	Displacement (ccm)	Power (hp)	Torque (Nm)	Engine cylinders
28736	suzuki	burgman 125	2014	Scooter	NaN	125.0	12.2	10.8	Sing cylinder
9366	gilera	nexus 300	2010	Scooter	3.7	292.0	NaN	NaN	Sing cylinder
30567	suzuki	tl 1000 r	1998	Sport	3.8	996.0	135.0	NaN	V
33825	vespa	gt60	2008	Scooter	3.1	244.0	21.7	20.2	Sing cylinder
12191	honda	cb 450 s	1990	Sport touring	3.4	447.0	44.0	NaN	Tw

5 rows × 28 columns



In [28]:

```
df.info()
df.describe()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 38472 entries, 0 to 38471
Data columns (total 28 columns):
 #   Column           Non-Null Count  Dtype  
 --- 
 0   Brand            38472 non-null   object  
 1   Model             38444 non-null   object  
 2   Year              38472 non-null   int64  
 3   Category          38472 non-null   object  
 4   Rating             21788 non-null   float64 
 5   Displacement (ccm) 37461 non-null   float64 
 6   Power (hp)         26110 non-null   float64 
 7   Torque (Nm)        16634 non-null   float64 
 8   Engine cylinder    38461 non-null   object  
 9   Engine stroke      38461 non-null   object  
 10  Gearbox            32675 non-null   object  
 11  Bore (mm)          28689 non-null   float64 
 12  Stroke (mm)        28689 non-null   object  
 13  Fuel capacity (lts) 31704 non-null   float64 
 14  Fuel system         27844 non-null   object  
 15  Fuel control        22008 non-null   object  
 16  Cooling system      34258 non-null   object  
 17  Transmission type   32861 non-null   object  
 18  Dry weight (kg)      22483 non-null   float64 
 19  Wheelbase (mm)       25493 non-null   float64 
 20  Seat height (mm)     24182 non-null   float64 
 21  Front brakes         36889 non-null   object  
 22  Rear brakes          36696 non-null   object  
 23  Front tire            31982 non-null   object  
 24  Rear tire             32008 non-null   object  
 25  Front suspension       26107 non-null   object  
 26  Rear suspension        25540 non-null   object  
 27  Color options          24327 non-null   object  
dtypes: float64(9), int64(1), object(18)
memory usage: 8.2+ MB
```

Out[28]:

	Year	Rating	Displacement (ccm)	Power (hp)	Torque (Nm)	Bore (mm)
<b>count</b>	38472.000000	21788.000000	37461.000000	26110.000000	16634.000000	28689.000000
<b>mean</b>	2003.195883	3.401574	552.515072	50.776040	64.527173	72.596000
<b>std</b>	20.083372	0.355631	545.394956	52.082094	63.884654	18.758000
<b>min</b>	1894.000000	1.400000	25.000000	0.300000	1.500000	1.000000
<b>25%</b>	2000.000000	3.200000	125.000000	12.000000	12.200000	57.000000
<b>50%</b>	2010.000000	3.400000	397.200000	30.000000	57.000000	73.000000
<b>75%</b>	2016.000000	3.700000	805.000000	77.000000	102.000000	88.000000
<b>max</b>	2022.000000	4.600000	8277.000000	804.000000	712.000000	176.000000



### 3. Datenbereinigung

```
In [29]: cols = [
    'Brand',
    'Year',
    'Category',
    'Displacement (ccm)',
    'Power (hp)',
    'Torque (Nm)',
    'Dry weight (kg)',
    'Seat height (mm)',
    'Fuel capacity (lts)',
    'Wheelbase (mm)'
]

df_clean = df[cols].dropna()
df_clean.sample(n=5)
```

Out[29]:

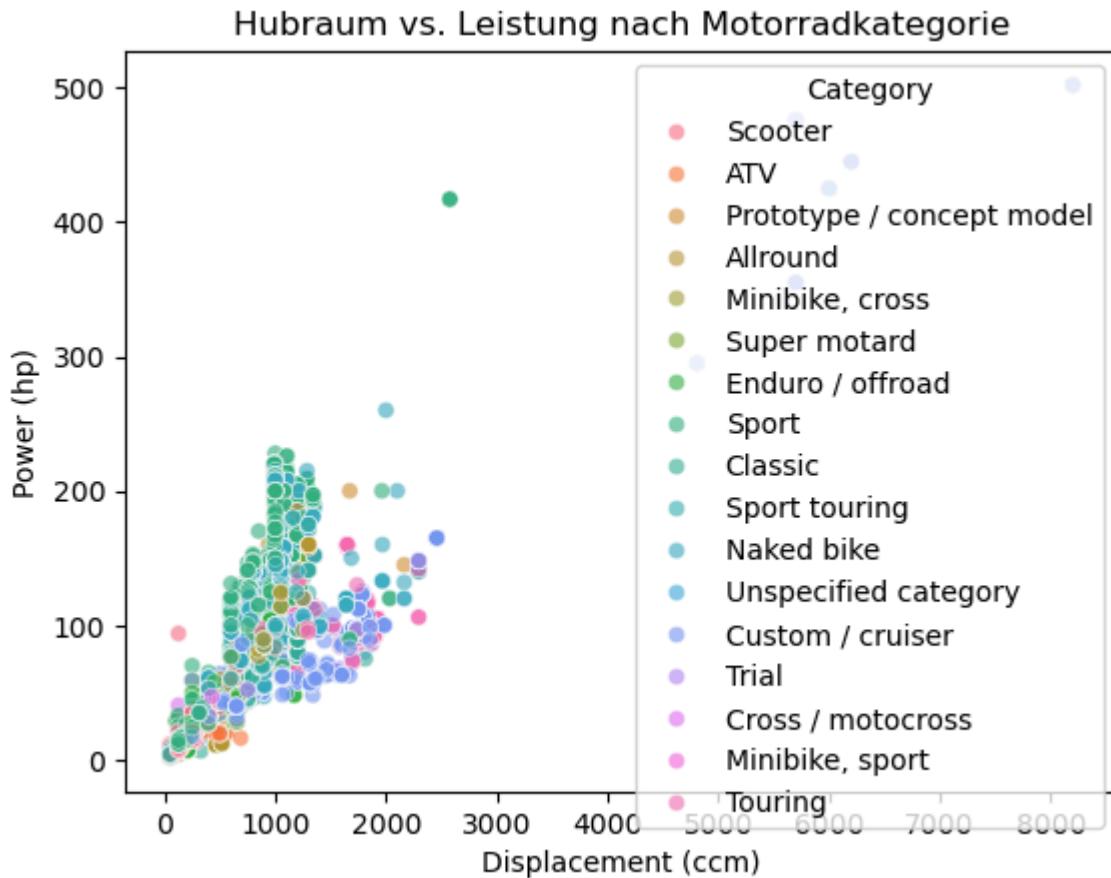
	Brand	Year	Category	Displacement (ccm)	Power (hp)	Torque (Nm)	Dry weight (kg)	Seat height (mm)	Fuel capacity (lts)
34492	wt motors	2020	Scooter	125.0	9.8	9.9	125.0	710.0	9.
23968	moto morini	2010	Naked bike	1187.0	140.0	123.0	196.0	830.0	18.
36804	yamaha	2007	Scooter	249.8	19.8	20.8	164.0	785.0	12.
17116	kawasaki	2012	Sport touring	1352.0	152.0	136.0	300.0	815.0	22.
6460	derbi	2011	Sport	49.9	8.5	5.9	110.0	810.0	13.

## 4. Leistungsanalyse

### 4.1 Hubraum vs. Leistung

```
In [30]: sns.scatterplot(
    data=df_clean,
    x='Displacement (ccm)',
    y='Power (hp)',
    hue='Category',
    alpha=0.6
)

plt.title("Hubraum vs. Leistung nach Motorradkategorie")
plt.show()
```

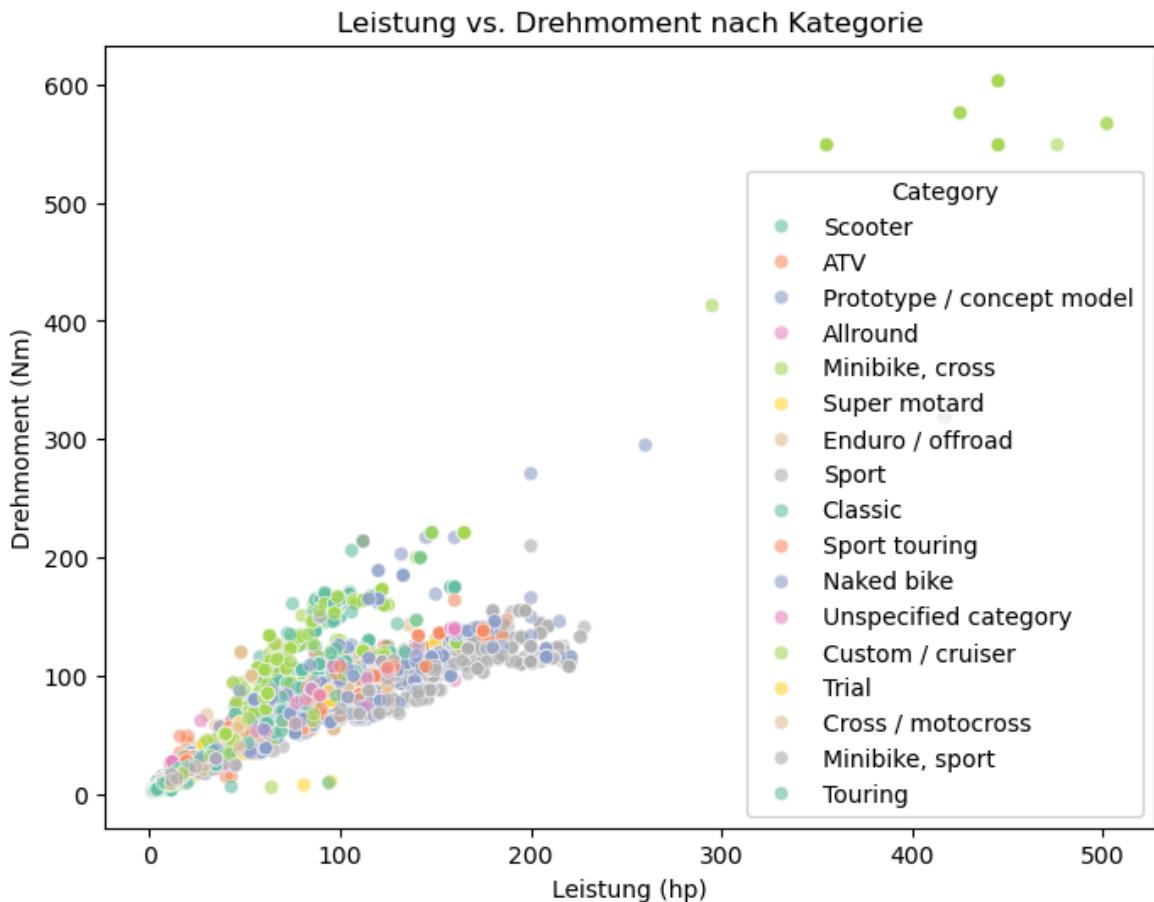


## 4.2 Leistung vs Drehmoment

```
In [31]: plt.figure(figsize=(8,6))

sns.scatterplot(
    data=df_clean,
    x='Power (hp)',
    y='Torque (Nm)',
    hue='Category',
    alpha=0.6,
    palette='Set2'
)

plt.title("Leistung vs. Drehmoment nach Kategorie")
plt.xlabel("Leistung (hp)")
plt.ylabel("Drehmoment (Nm)")
plt.show()
```



### 4.3 Hubraum & Drehmoment nach Kategorie

```
In [32]: plt.figure(figsize=(10,6))

sns.boxplot(
    data=df_clean,
    x='Category',
    y='Displacement (ccm)',
    hue='Category',
    palette='Set2',
    legend=False
)

plt.title("Hubraumverteilung nach Motorradkategorie")
plt.xlabel("Kategorie")
plt.ylabel("Hubraum (ccm)")
plt.xticks(rotation=45)
plt.show()

plt.figure(figsize=(10,6))

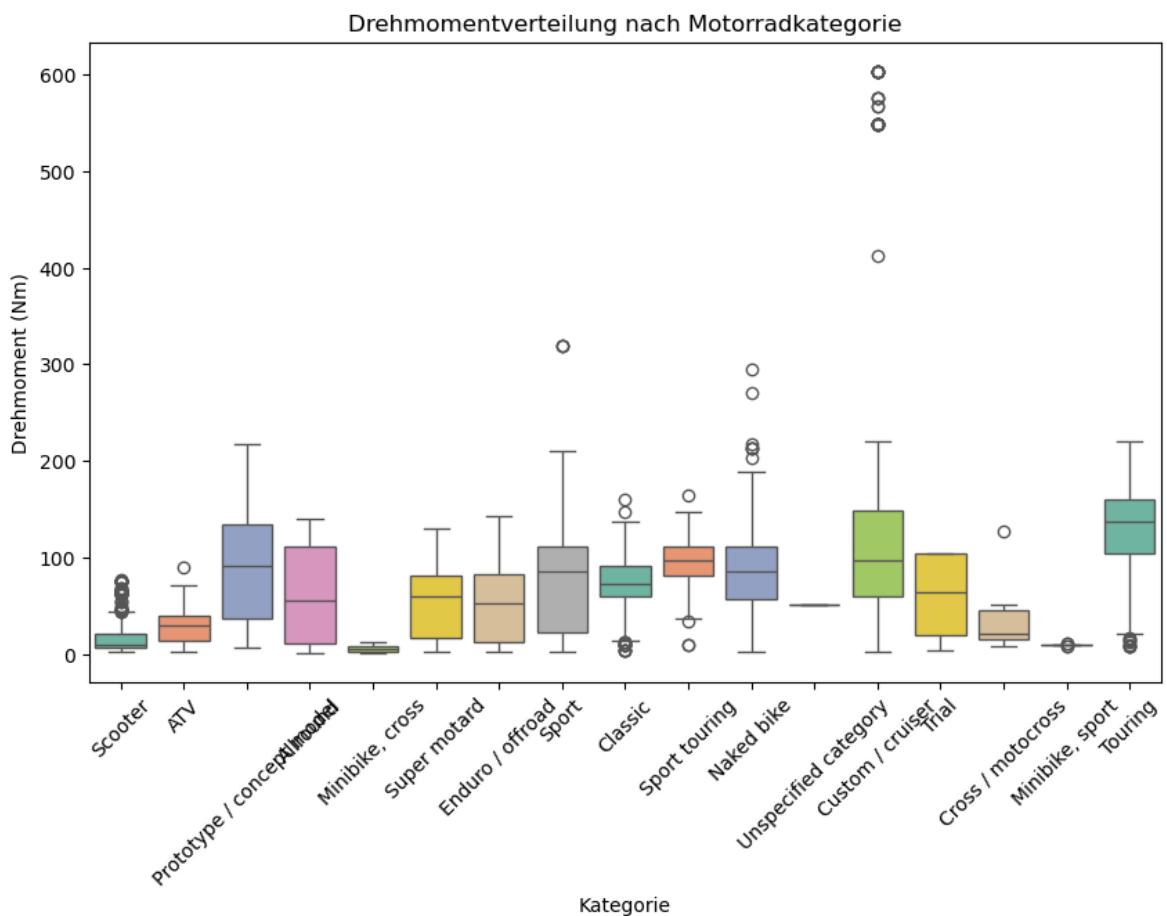
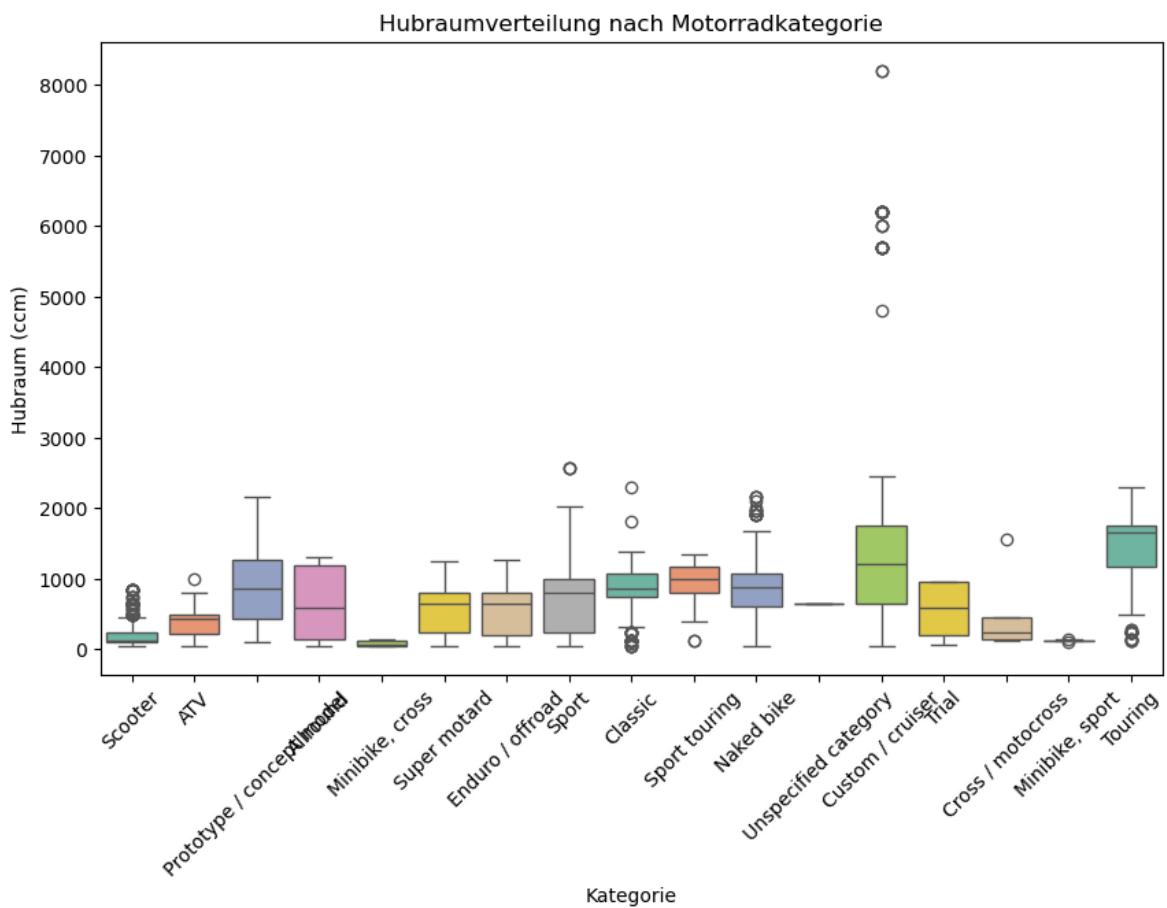
sns.boxplot(
    data=df_clean,
    x='Category',
    y='Torque (Nm)',
    hue='Category',
    palette='Set2',
    legend=False
)

plt.title("Drehmomentverteilung nach Motorradkategorie")
```

```

plt.xlabel("Kategorie")
plt.ylabel("Drehmoment (Nm)")
plt.xticks(rotation=45)
plt.show()

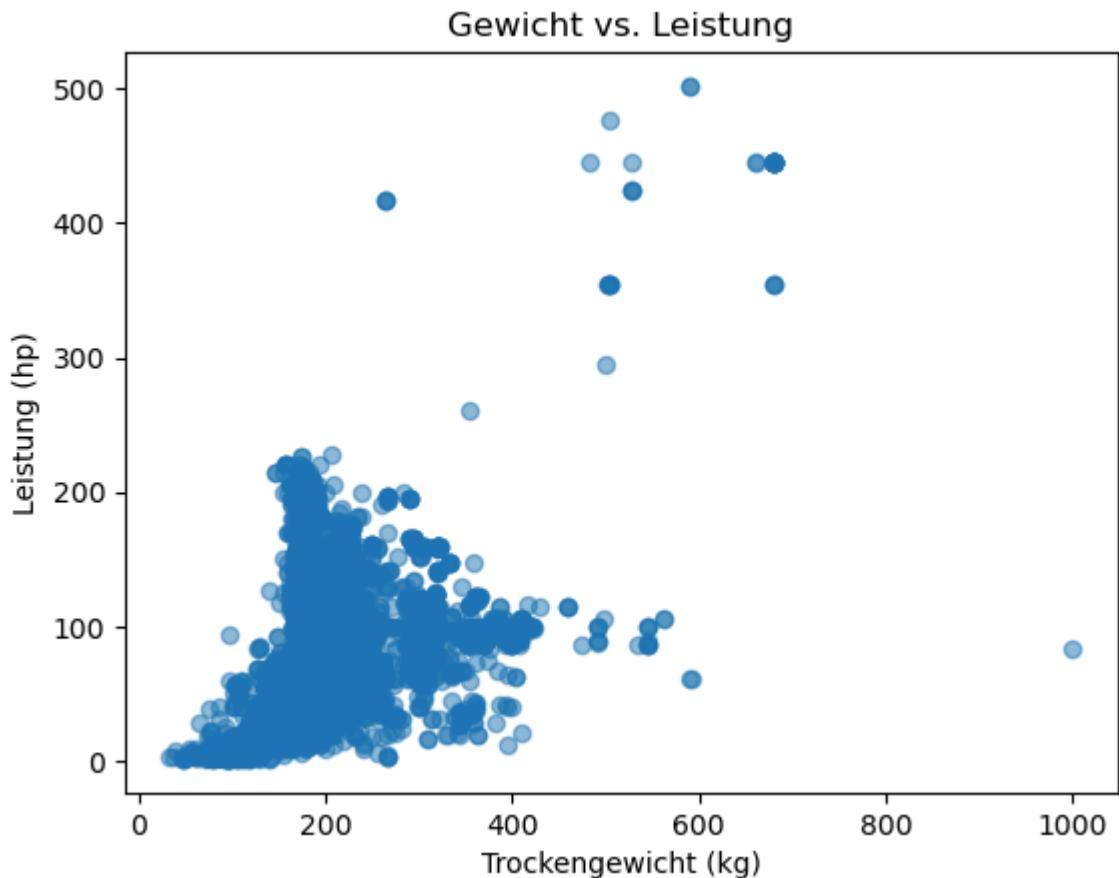
```



## 4.4 Gewicht vs. Leistung

```
In [33]: plt.scatter(
    df_clean['Dry weight (kg)'],
    df_clean['Power (hp)'],
    alpha=0.5
)

plt.xlabel("Trockengewicht (kg)")
plt.ylabel("Leistung (hp)")
plt.title("Gewicht vs. Leistung")
plt.show()
```

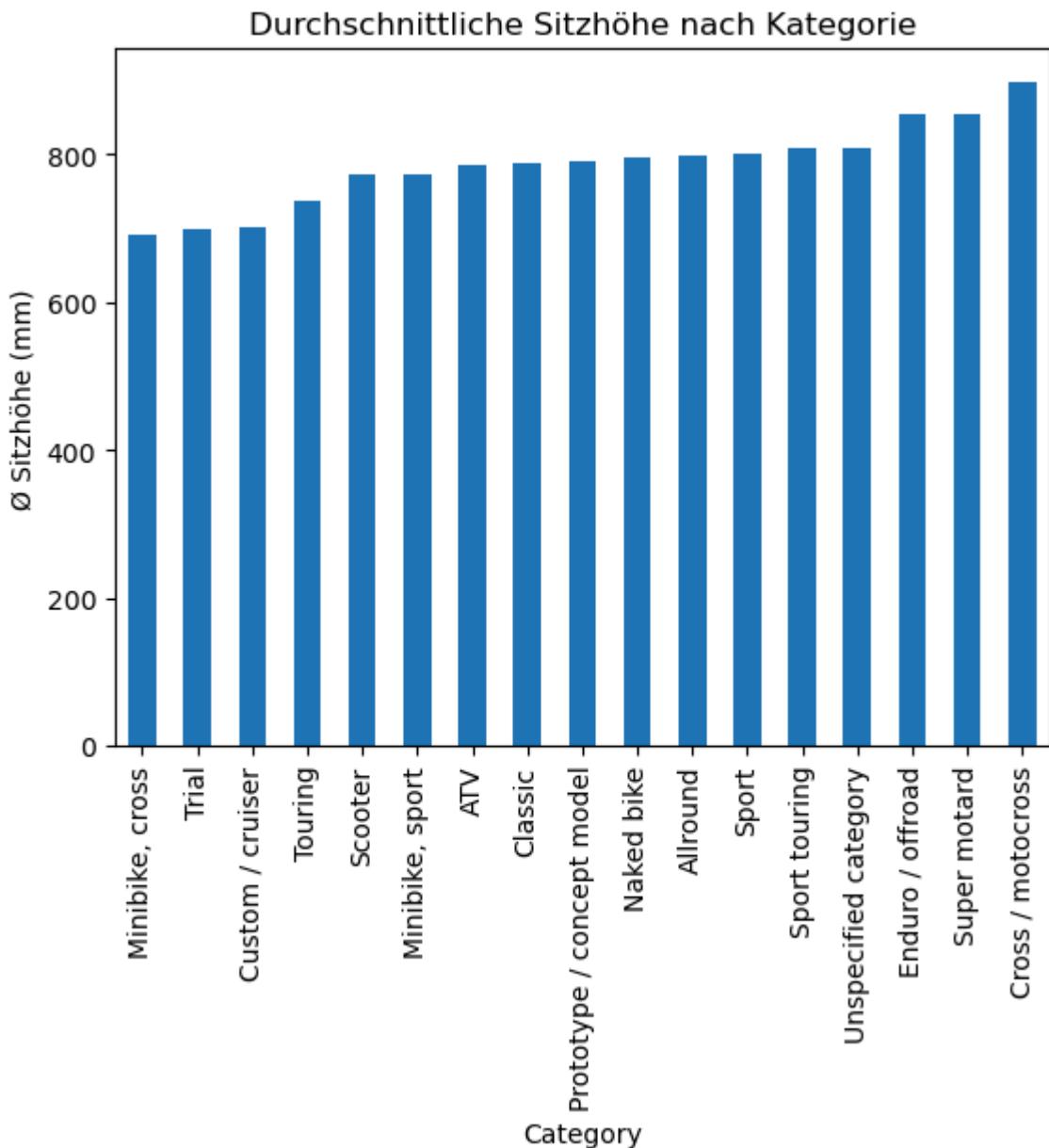


## 5. Designanalyse

### 5.1 Sitzhöhe nach Kategorie

```
In [34]: df_clean.groupby('Category')[ 'Seat height (mm)' ] \
    .mean() \
    .sort_values() \
    .plot(kind='bar')

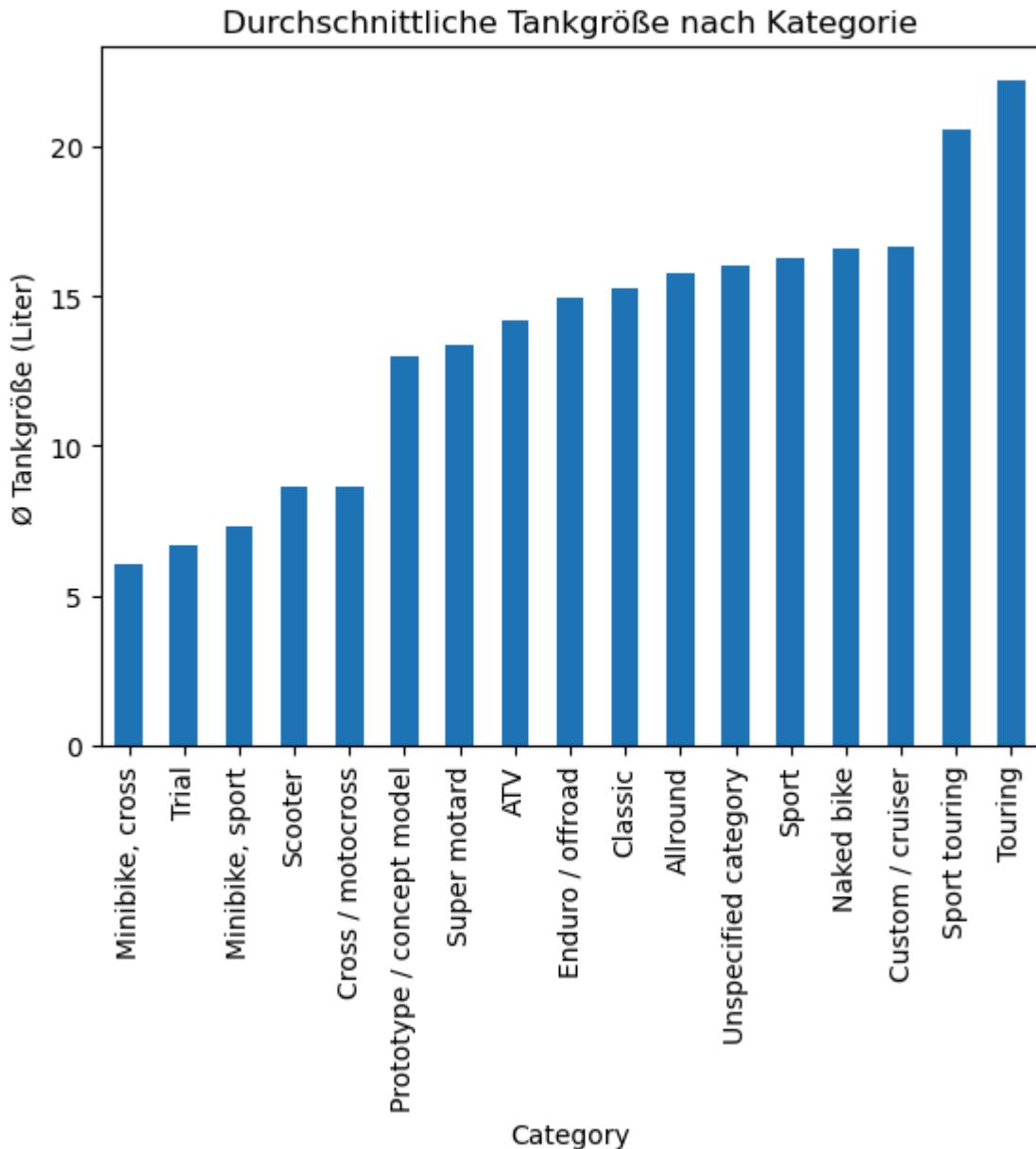
plt.ylabel("Ø Sitzhöhe (mm)")
plt.title("Durchschnittliche Sitzhöhe nach Kategorie")
plt.show()
```



## 5.2 Tankgröße nach Kategorie

```
In [35]: df_clean.groupby('Category')['Fuel capacity (lts)'] \
    .mean() \
    .sort_values() \
    .plot(kind='bar')

plt.ylabel("Ø Tankgröße (Liter)")
plt.title("Durchschnittliche Tankgröße nach Kategorie")
plt.show()
```



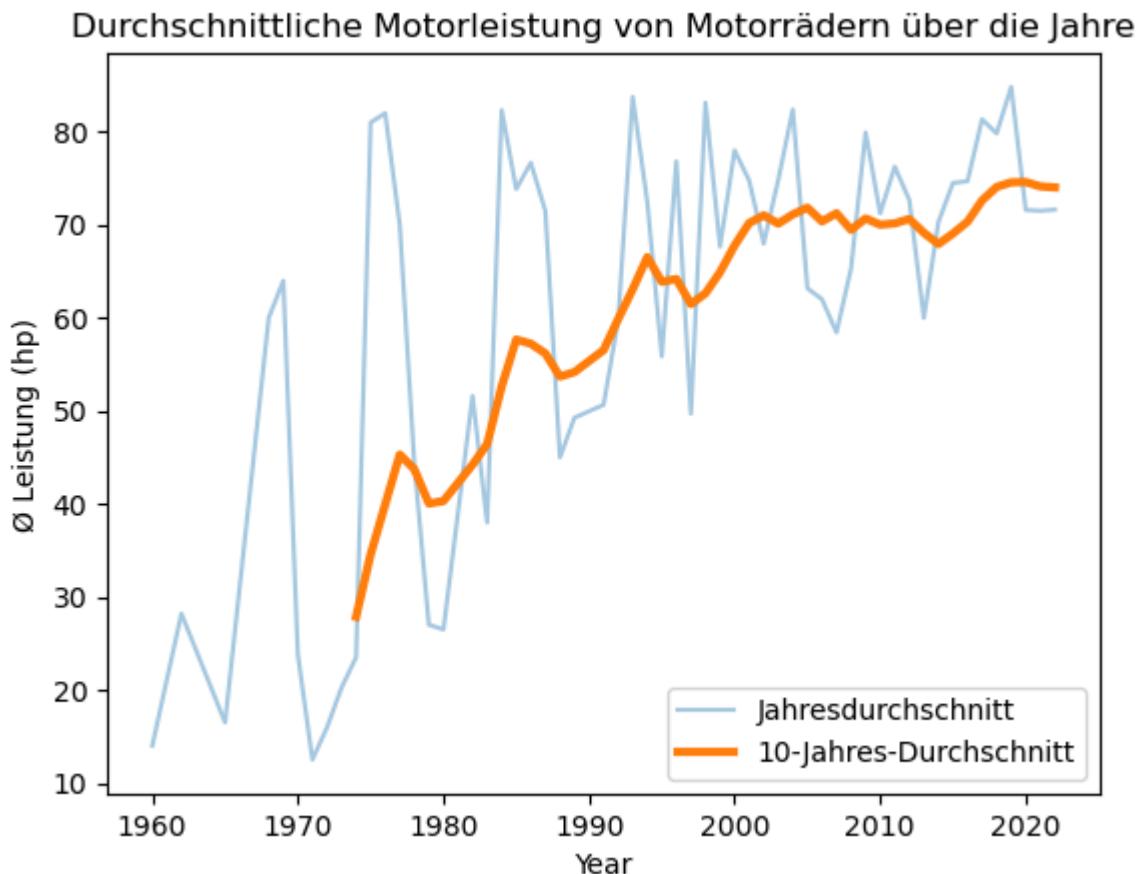
## 6. Zeitliche Entwicklung

```
In [36]: avg_power_per_year = df_clean.groupby('Year')[['Power (hp)']].mean()

avg_power_per_year.plot(label='Jahresdurchschnitt', alpha=0.4)

avg_power_per_year.rolling(window=10).mean().plot(
    label='10-Jahres-Durchschnitt',
    linewidth=3
)

plt.ylabel("Ø Leistung (hp)")
plt.title("Durchschnittliche Motorleistung von Motorrädern über die Jahre")
plt.legend()
plt.show()
```



## 7. Fazit

- Man sieht einen klaren Anstieg der Leistung über die Jahre. Die Leistung hängt einerseits vom Hubraum als auch vom Gewicht des Motorrads ab.
- Das Design der Motorräder hängt von den Motorradkategorien ab, da diese für unterschiedliche Zwecke gebaut werden.