

# Analyse Exploratoire des Données de Vent (Mensuelles par Commune)

## 1. ANALYSES SIMPLES VISION GLOBALE

Q1 : Quels sont les ordres de grandeur des variables ?

```
df.describe()
```

But : détecter les valeurs aberrantes, calibrer les échelles, connaître les moyennes.

Q2 : Comment se répartissent les valeurs sur l'ensemble des communes/mois ?

```
sns.histplot(df['gust_max'], bins=30, kde=True)
```

But : voir si la distribution est normale, asymétrique ou extrême (long tail).

## 2. ANALYSES TEMPORELLES (FRANCE ENTIÈRE)

Q3 : Quelle est la tendance mensuelle globale du vent ?

```
df.groupby('date_mois')[['wind_speed_10m_mean', 'gust_max']].mean().plot(marker='o')
```

But : identifier les pics anormaux, les mois calmes vs à risque.

Q4 : Y a-t-il une saisonnalité du vent ?

```
df['mois'] = pd.to_datetime(df['date_mois']).dt.month
```

```
df.groupby('mois')[['gust_max', 'wind_speed_100m_mean']].mean().plot()
```

But : exploiter des effets de mois (variables saisonnières) en modélisation.

## 3. ANALYSES SPATIALES (COMMUNES)

Q5 : Quelles communes sont les plus exposées en rafale max ?

```
df.groupby('code_commune')['gust_max'].mean().sort_values(ascending=False).head(20)
```

But : identifier les zones structurellement à risque.

Q6 : Ces communes sont-elles géographiquement cohérentes ?

En affichant sur une carte (si tu as les coordonnées ou un shapefile).

Q7 : Y a-t-il des communes avec haute variance intermensuelle ?

```
df.groupby('code_commune')['wind_speed_10m_mean'].std().sort_values(ascending=False).head(20)
```

But : détecter l'instabilité structurelle.

## 4. INTERACTIONS SPATIO-TEMPORELLES

Q8 : Évolution du vent pour une commune précise sur plusieurs mois

```
df[df['code_commune'] == 'XXXXXX'].set_index('date_mois')[['wind_speed_10m_mean', 'gust_max']].plot()
```

But : préparer le terrain pour des séries temporelles locales (contrats géolocalisés).

Q9 : Détection des événements extrêmes (rafale > seuil)

```
df['gust_extreme'] = df['gust_max'] > 30
```

```
df.groupby('date_mois')['gust_extreme'].sum().plot(kind='bar')
```

But : observer quand ont eu lieu des pics météo dangereux.

## 5. ANALYSES MULTIVARIÉES & SEGMENTATION

Q10 : Quelles sont les corrélations entre variables météo ?

```
sns.heatmap(df.drop(columns=['date_mois', 'code_commune']).corr(), annot=True)
```

But : réduire la redondance pour la modélisation.

Q11 : Peut-on segmenter les communes en fonction de leur profil météo ?

```
from sklearn.cluster import KMeans
```

```
X = df.groupby('code_commune')[['gust_max', 'wind_speed_10m_mean']].mean()
```

```
X['cluster'] = KMeans(n_clusters=4, random_state=42).fit_predict(X)
```

But : enrichir ta base de modélisation avec une typologie météo.