



Trouve ton job dans l'IA

Fengfeng ZHANG
Asmae ADNADANE
David KOROBETSKI

Formation D  v IA2

03-02-2023

Sommaire

01 Contexte et Objectifs du projet

| | | |
|--------------------------------|-----------|---|
| □ Objectifs | | 3 |
| □ Mission de travail en groupe | | 5 |

02 Méthodologie

| | | |
|---|-----------|---|
| □ Intégration / Préparation / Nettoyage | | 7 |
| □ Analyse descriptive et exploratoire | | 7 |
| □ Modélisation | | 8 |
| □ Développement d'une application web | | 8 |

03 Résultats

| | | |
|---|-----------|----|
| □ Situations actuelles de recrutement sur l'IA | | 10 |
| □ Prévoyez votre futur paiement d'emploi avec notre application | | 22 |

04 Conclusion





FIND JOBS



Compétences

SQL
HTML
...

où

\$350

Intitulé
Data analyst
Developer
Data scientist
...



[Lancer la recherche](#) [Avis sur les entreprises](#) [Estimation de salaire](#) [Télécharger votre CV](#) [Connexion](#) [Entreprises / Publier une annonce](#)

Quoi AI

Q

Où Ville ou code postal

📍

Rechercher

Astuce : indiquez une ville ou un code postal dans la barre "où" afin d'afficher des résultats localisés.

Date de publication

Posté par

Télétravail

Estimation du salaire

Type de poste

Secteurs

Lieu

Entreprise

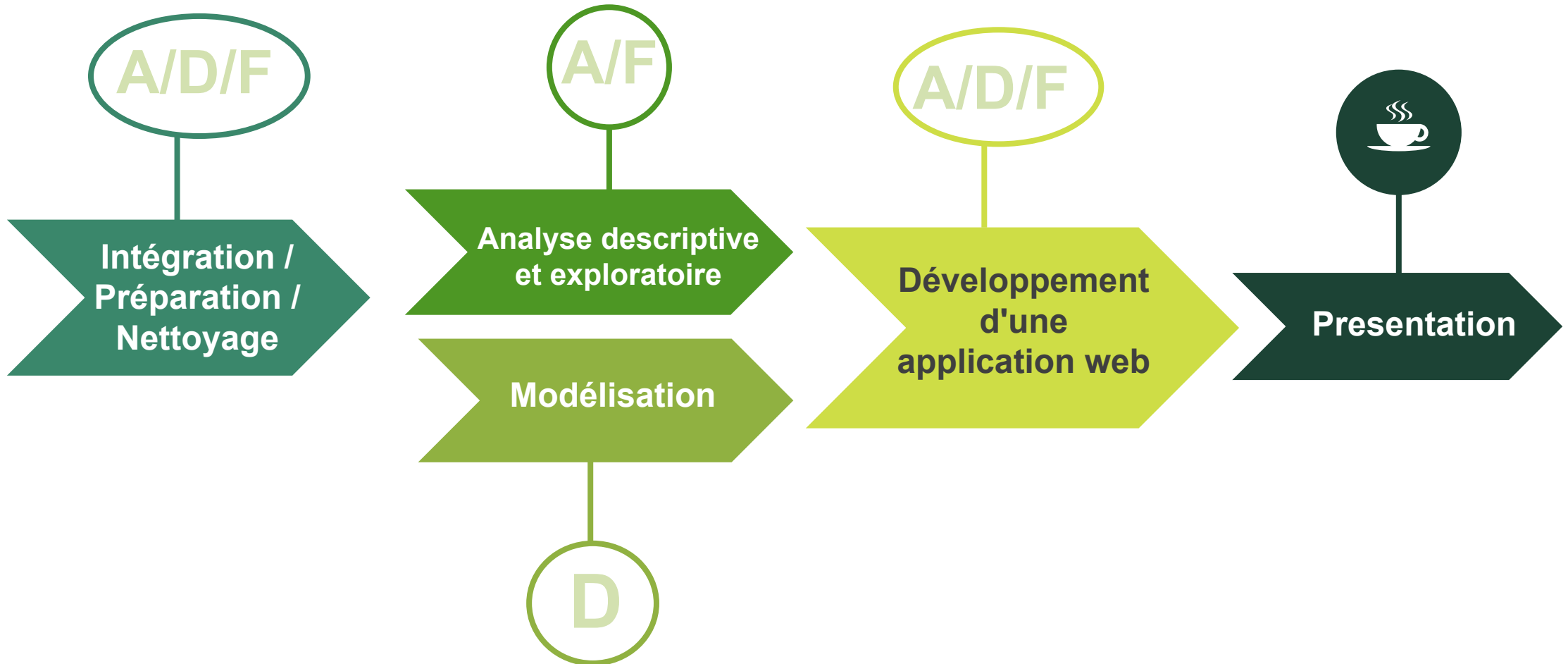
Objectifs du projet

Analyser des données d'offres d'emploi en rapport avec votre futur métier dans l'IA

Faire un meilleur usage des informations de recrutement publiées sur Internet



Mission de travail en groupe



T

Trouve ton job
Free

Boards

Members

Settings

Workspace views

Table

Calendar

Your boards

Planning

Planning

Workspace visible

Board

infos

Ordre de priorité

+ Add a card

To do

Première version du diapo

+ Add a card

in progress

Créer le fichier de présentation sur GD et le partager avec le groupe

Développer l'application

+ Add a card

test

Analyse Exploratoire

+ Add a card

done

Nettoyer le dataset

Modèle de machine learning

Formater et imputer la colonne salaire

Nettoyer la colonne intitulé de poste

Formater type de contrat

Créer un planning trello

Créer un dépôt GitHub

+ Add a card

Méthodologie

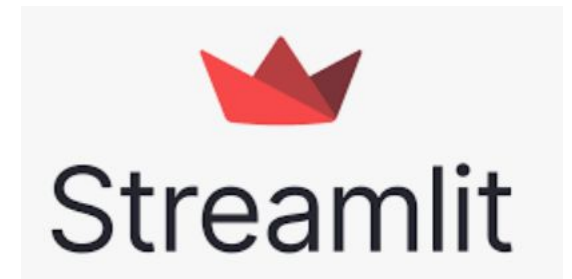
Intégration / Préparation / Nettoyage

**Analyse descriptive et
exploratoire**

python plotly

Modélisation

**Développement d'une
application web**





Modélisation

```
# Pipeline

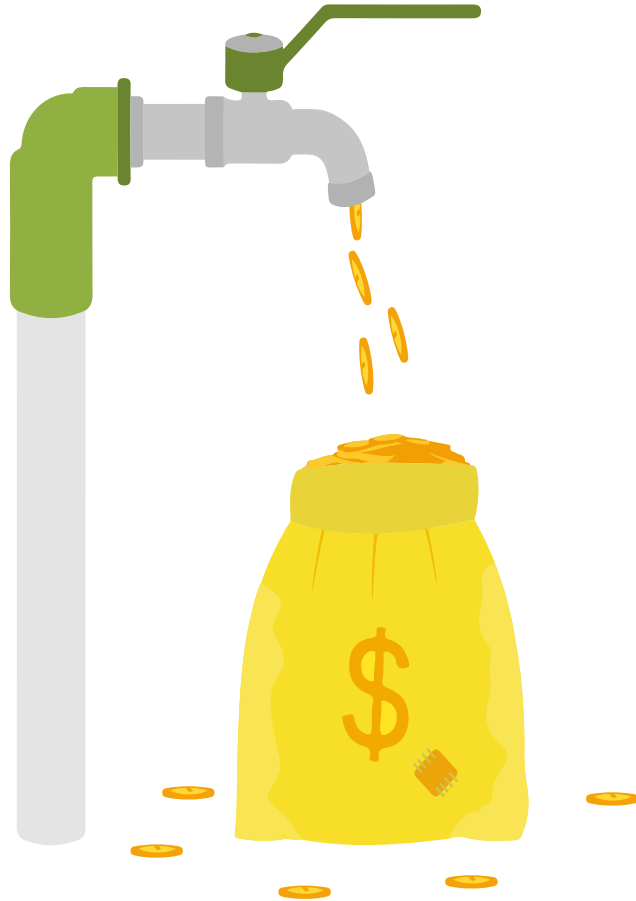
# dictionnaire utilisé pour la recherche des meilleurs de paramètres
...
params_RFR = {
    "criterion" : ['absolute_error', 'friedman_mse'],
    "random_state" : [n for n in range(39, 67)],
    "n_estimators" : [n for n in range(8, 27)]
}
...

pipe_cat = Pipeline(
    steps=[
        ('pipe_imp', SimpleImputer(strategy='most_frequent')),
        ('pipe_enc', OneHotEncoder(sparse=False))
    ]
)

tf_cat = ColumnTransformer(
    transformers=[
        ('tf_cat', pipe_cat, ['Intitulé du poste', 'Nom de la société', 'Type de contrat']),
        ('tf_comp', CountVectorizer(), 'competences')
    ]
)

RFR_pipe_max = Pipeline(
    steps=[
        ('transformation', tf_cat),
        ('model', RandomForestRegressor(n_estimators=26, random_state=66, criterion='absolute_error'))
    ]
)

RFR_pipe_min = Pipeline(
    steps=[
        ('transformation', tf_cat),
        ('model', RandomForestRegressor(n_estimators=9, random_state=40, criterion='friedman_mse'))
    ]
)
```

Evaluation du score de prédiction pour le salaire maximum

```
RFR_pipe_max.fit(X_train, y_train)
y_max_pred = RFR_pipe_max.predict(X_test)
print("RFR:", round(r2_score(y_test, y_max_pred), 5))
# print("best params : ", RFR_pipe_max['model'].best_params_)
```

RFR: 0.86616

Evaluation du score de prédiction pour le salaire minimum

```
RFR_pipe_min.fit(X_train, y_train)
y_min_pred = RFR_pipe_min.predict(X_test)
print("RFR:", round(r2_score(y_test, y_min_pred), 5))
# print("best params : ", RFR_pipe['model'].best_params_)
```

RFR: 0.83544

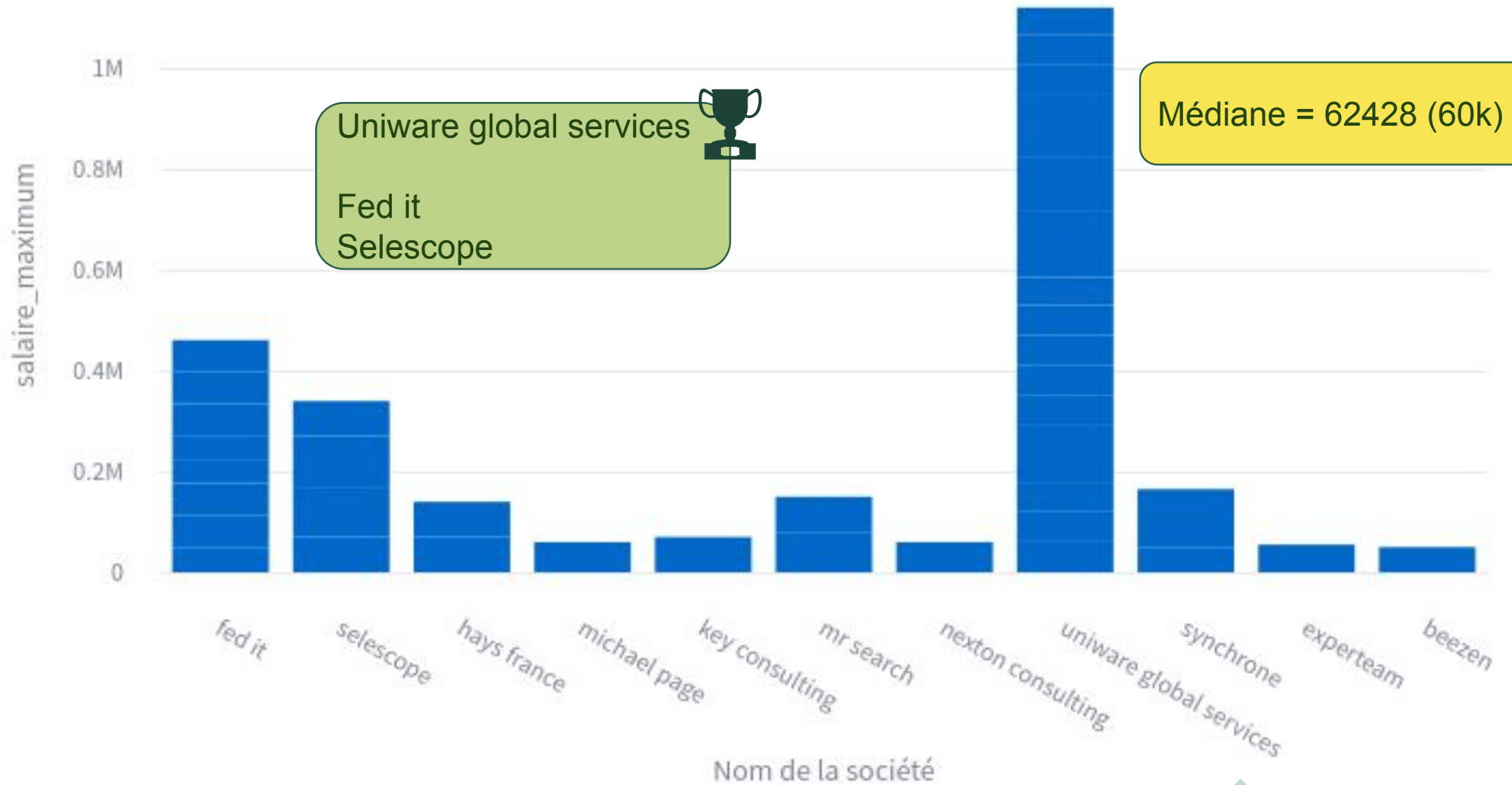
Résultats

Situations actuelles de
recrutement sur l'IA



Salaire maximal

Masse salariale en fonction de la colonne 'Nom de la société'



Salaire maximal

Masse salariale en fonction de la colonne 'Lieu'



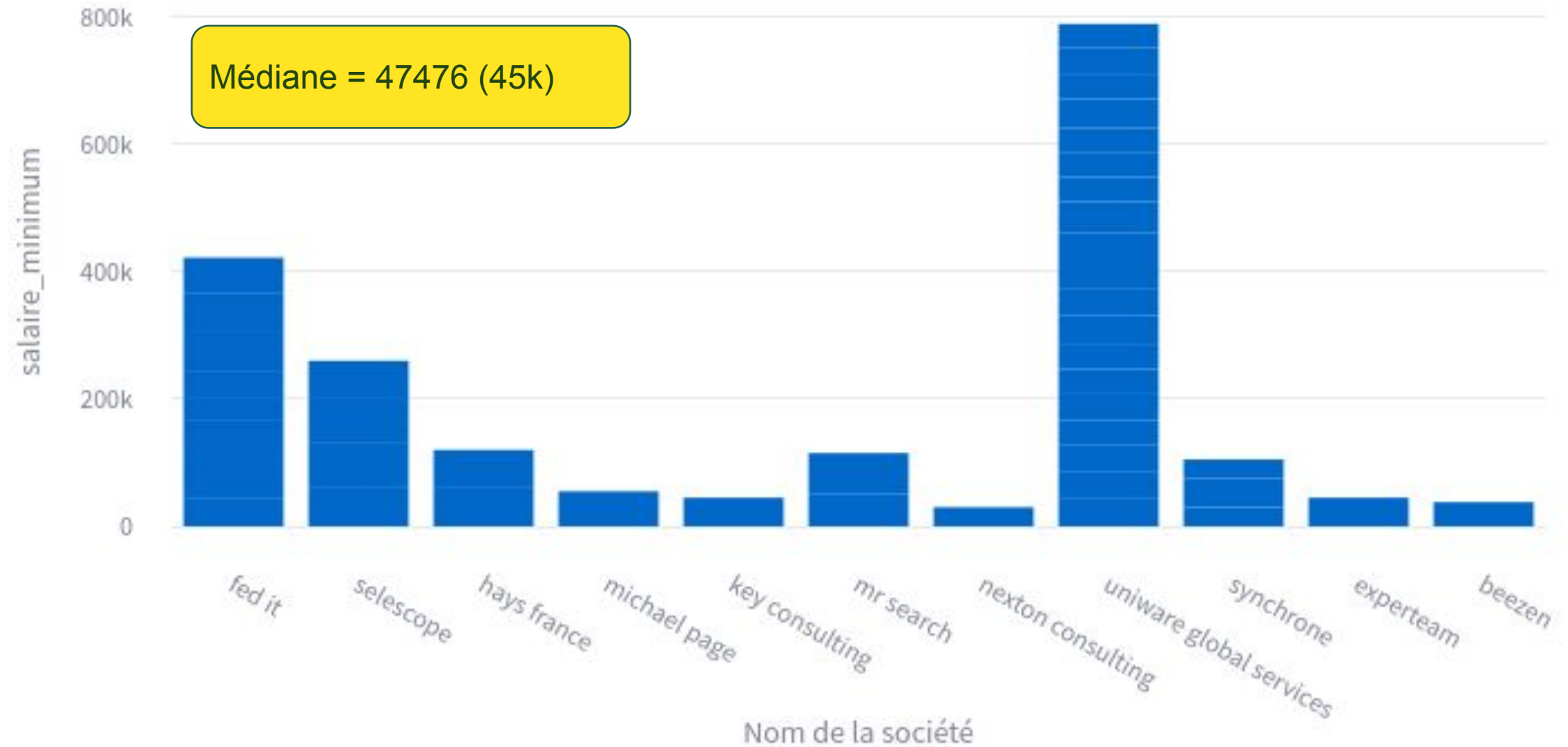
Salaire maximal

Masse salariale en fonction de la colonne 'Intitulé du poste'



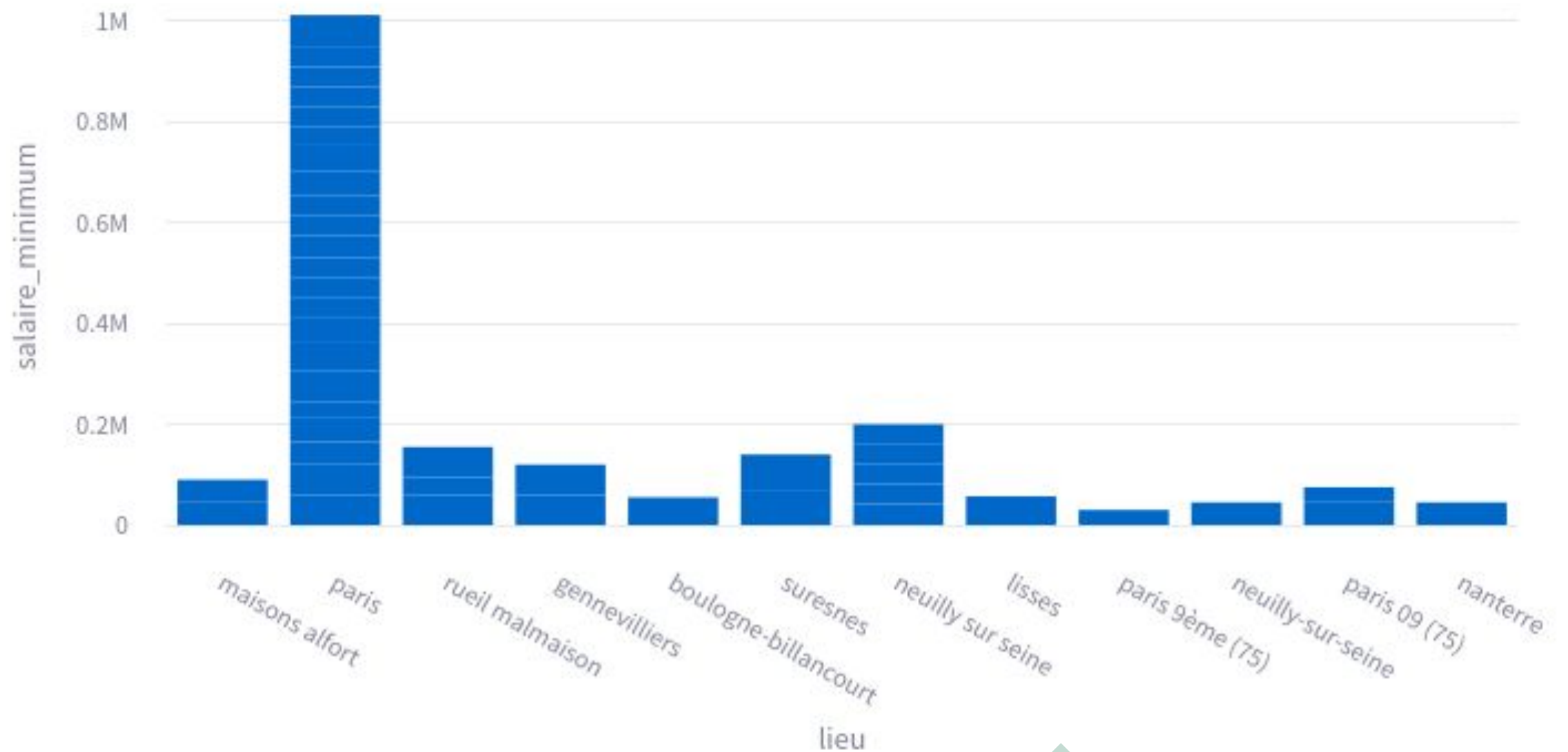
Salaire minimum

Masse salariale en fonction de la colonne 'Nom de la société'



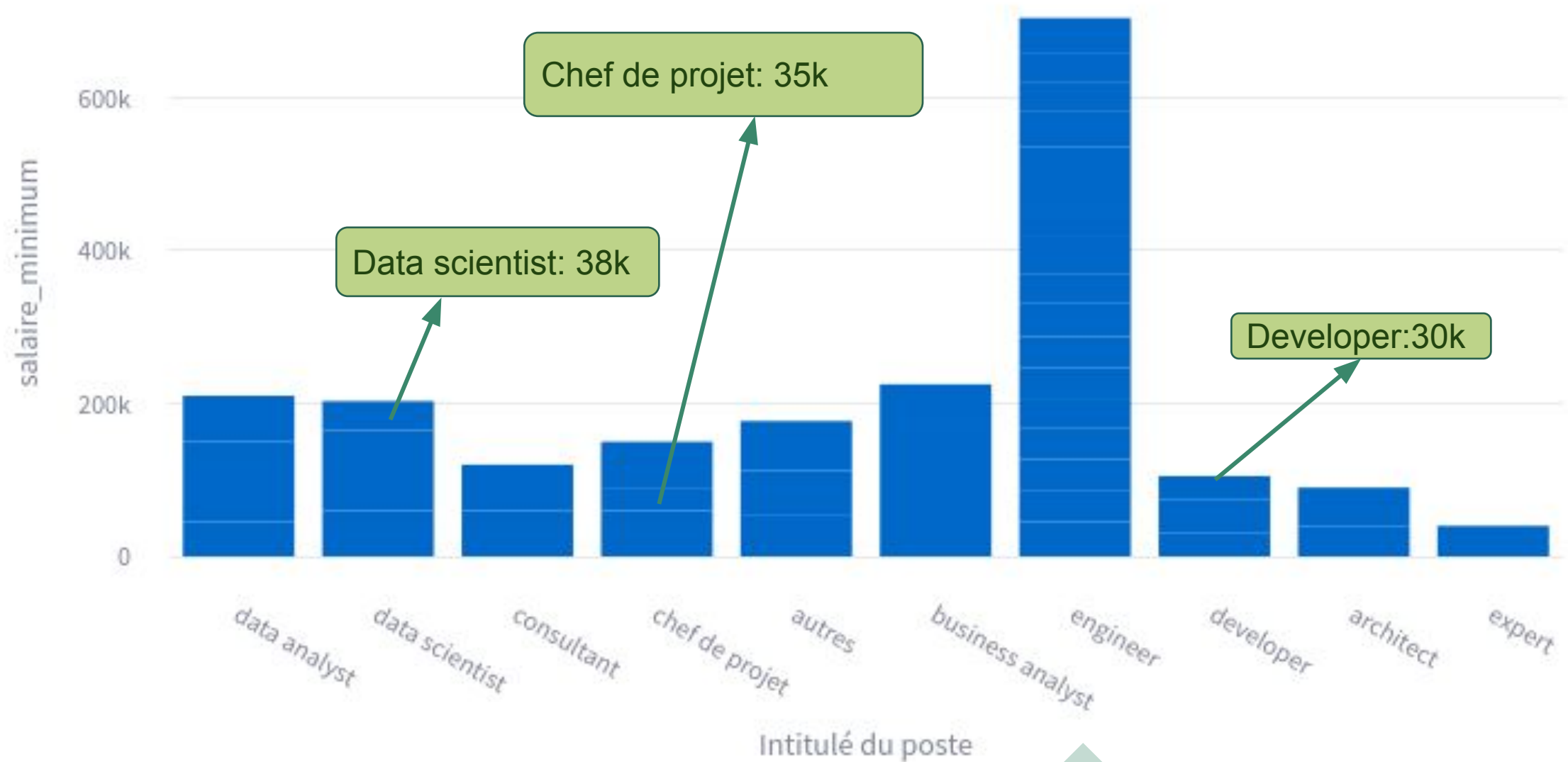
Salaire minimum

Masse salariale en fonction de la colonne 'Lieu'



Salaire minimum

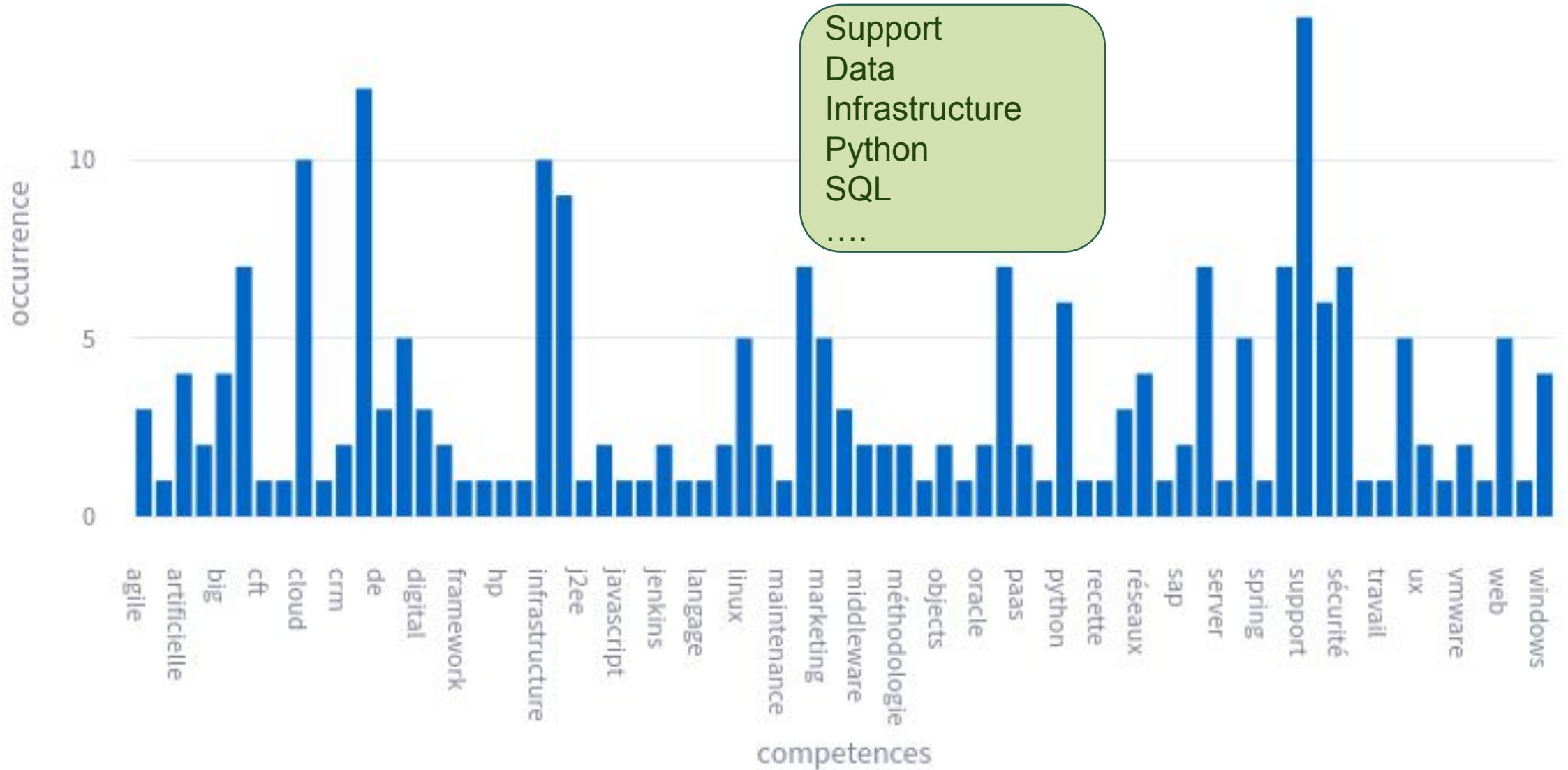
Masse salariale en fonction de la colonne 'Intitulé du poste'



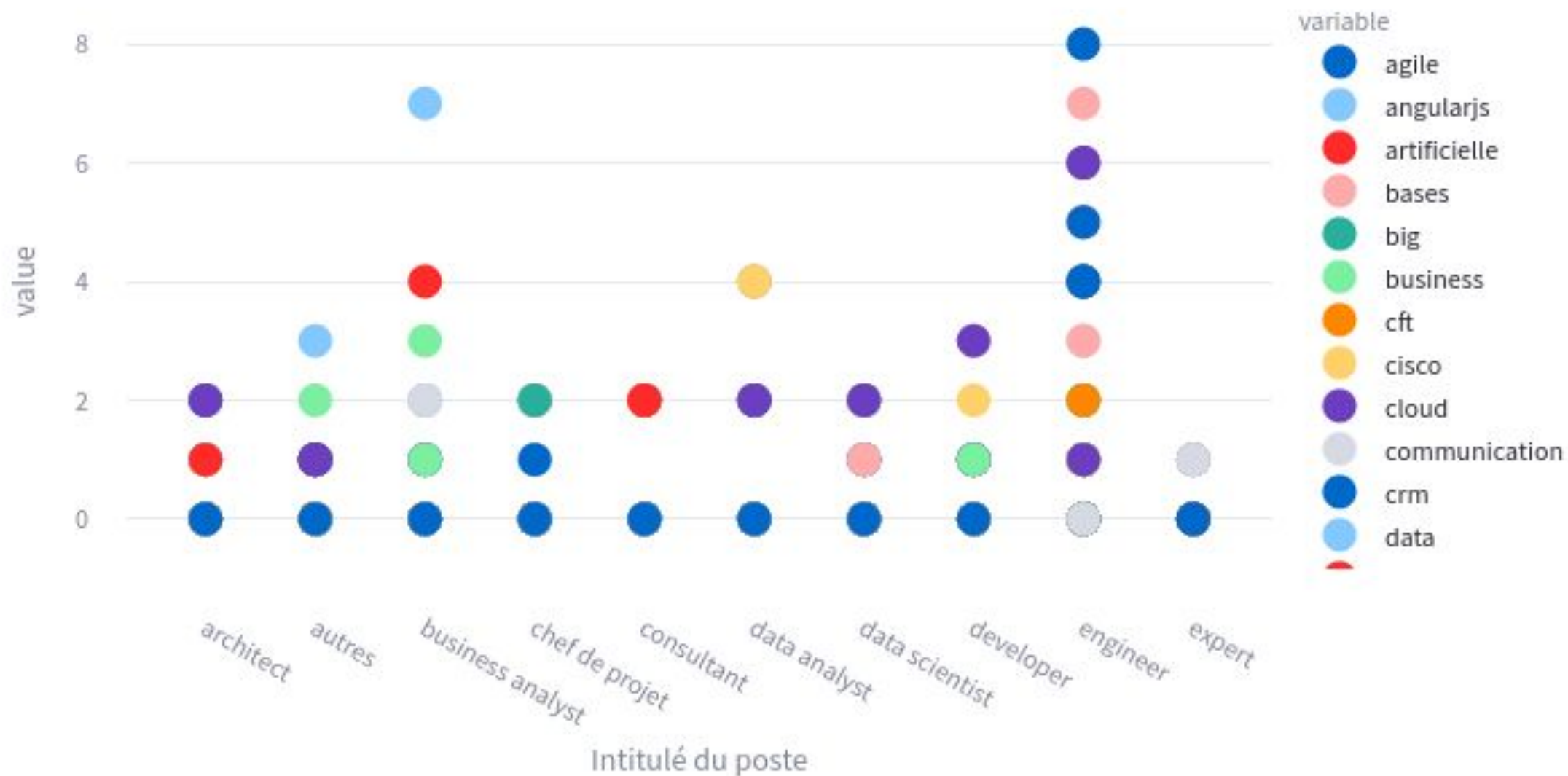


**Les compétences les plus
recherchées**

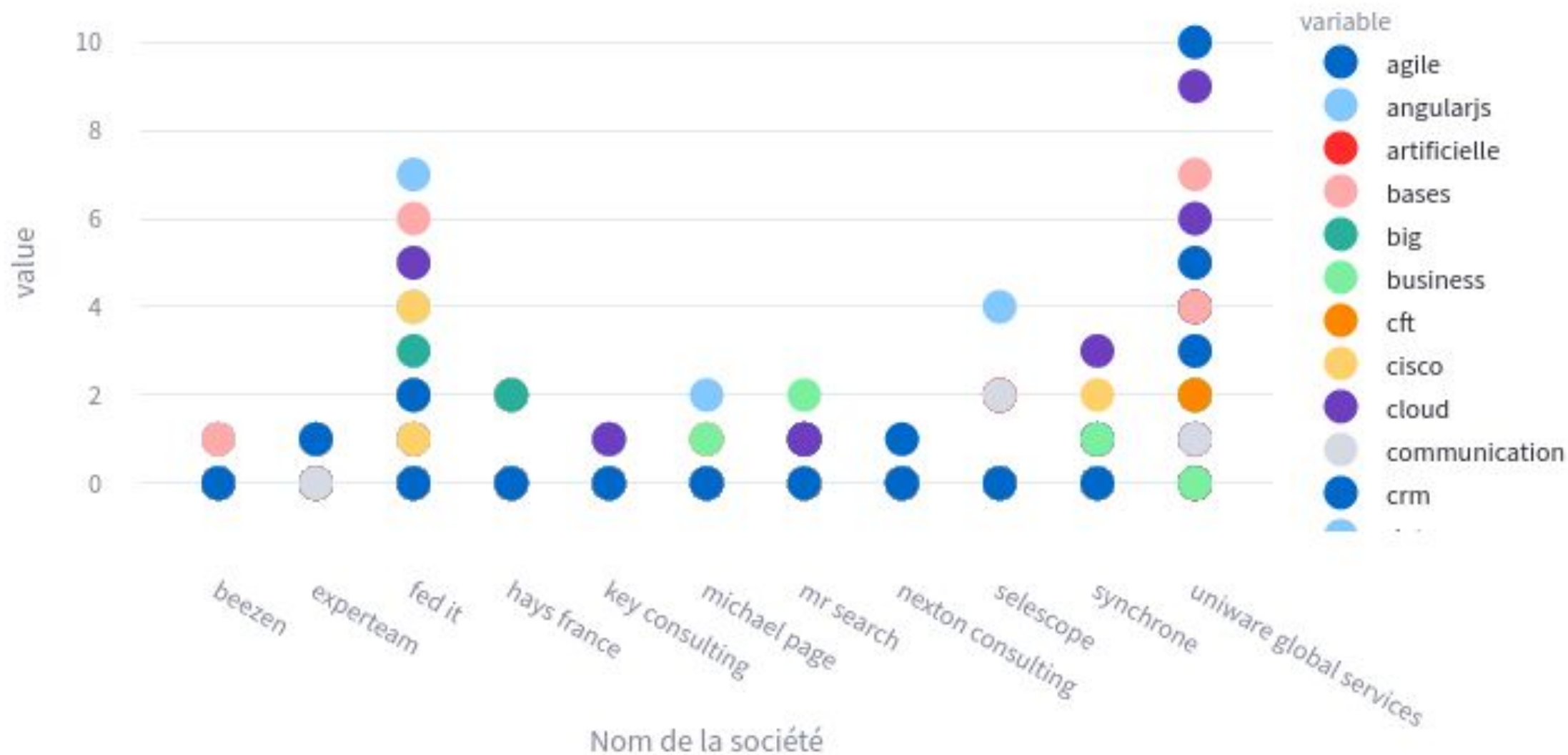
Les compétences les plus demandées



Répartition des compétences en fonction de la colonne 'Intitulé du poste'



Répartition des compétences en fonction de la colonne 'Nom de la société'



Analyse descriptive et exploratoire



FIND JOB

Prévoyez votre futur
salaire avec notre
application web





Conclusion

01

Partie résultats

- ❖ Feedback
- ❖ UI
- ❖ Scalabilité
- ❖ Déploiement

02

Partie sur la technique

- ❖ Données manquantes
- ❖ r^2 score correct
- ❖ Bon travail dans la groupe
- ❖ Nettoyage approximatif
- ❖ Exploration à approfondir





THANK YOU