

Analyse de données marketing

LAB DM - Rida MESSAOUDENE

OLIVIER VASTEL
NAISSIANE LABROSSE
MARGOT HERAIL

DC5C

1 - Définition de la Problématique

Identifiez une question de recherche ou un objectif d'analyse spécifique que vous souhaitez explorer à travers les datasets fournis.



Qu'elles sont les **tendances du marché du travail** dans les secteurs du **marketing**, de la **data** et de **l'IT** en terme de **rémunération** ?

Data set choisis :

- Données sur les **salaires** par **titre de poste** et par **pays**.

Ce dataset est composé des colonnes suivantes :

L'âge, le genre, le niveau d'études, le titre de poste, les années d'expériences, le salaire, le pays et l'ethnie.

En analysant le dataset, nous avons remarqué qu'ils y avaient une prépondérance de titres de poste liés au secteur du **marketing**, de la **data** et de **l'IT**.

Nous avons alors décidé de supprimer les autres postes afin de prédire les revenus de ces métiers en particulier, ce qui nous fait un jeu de données composé de **6 698 lignes**.



ALLER SUR NOTRE GITHUB

2 - Construction des hypothèses

Formulez des hypothèses basées sur votre compréhension des enjeux marketing et clientèle, que vous testerez par la suite grâce à l'analyse de données.

HYPOTHÈSES

01

Plus on a **d'expériences**, plus notre **salaire** sera **élevé**.

02

Plus on est **âgé**, plus notre **salaire** sera **élevé**.

03

Plus on a fait de **longues études**, plus notre **salaire** sera **élevé**.

04

Les salaires **diffèrent** d'un **pays** à l'autre.

05

Les salaires **diffèrent** d'une **ethnie** à l'autre.

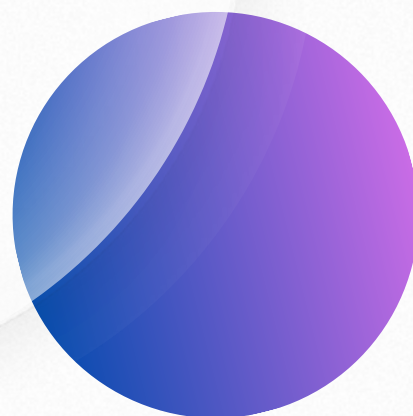
3 - Modélisation prédictive

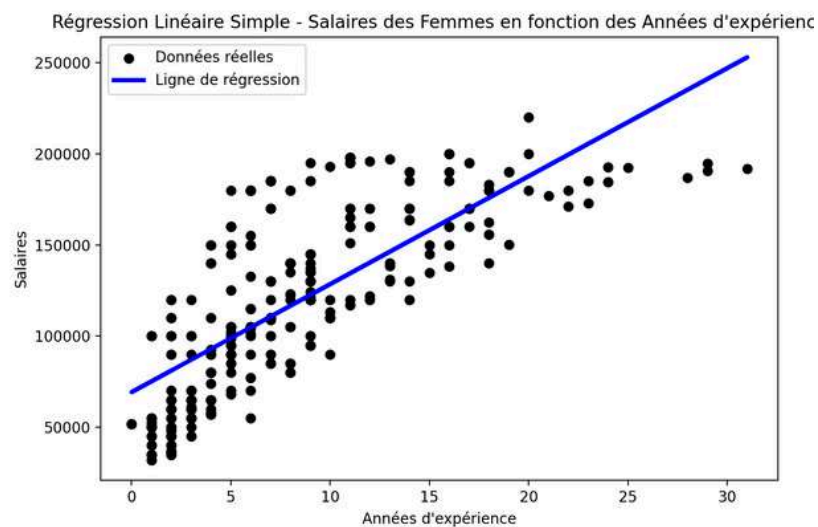
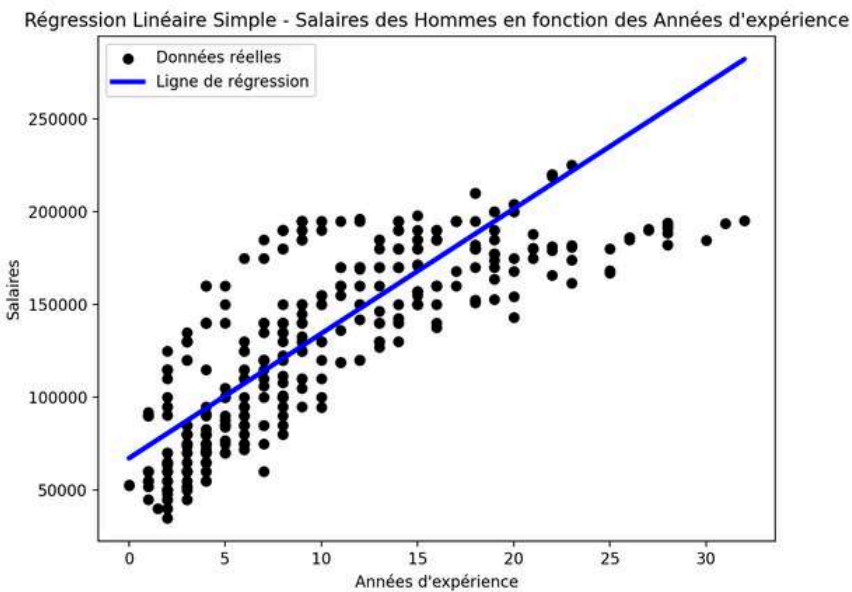
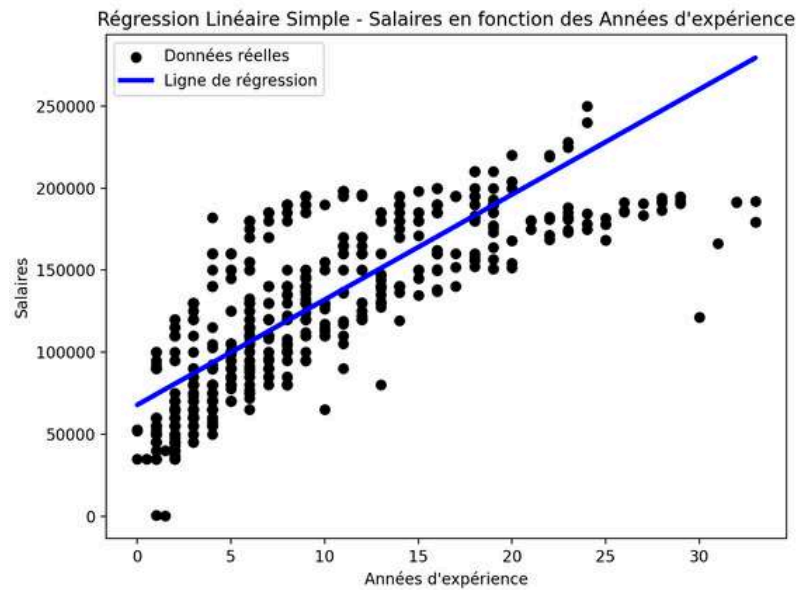
Selon votre problématique, appliquez des techniques de machine learning pour prédire des résultats ou classifier des données, en justifiant vos choix de modèles.

ÉTAPE 1:



Analyse de **l'impact des années d'expériences** sur les **salaires**.



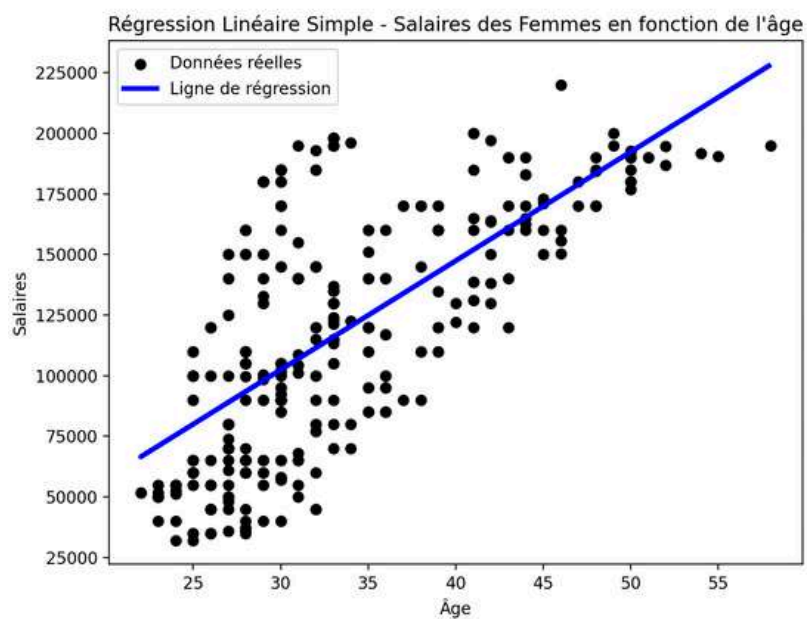
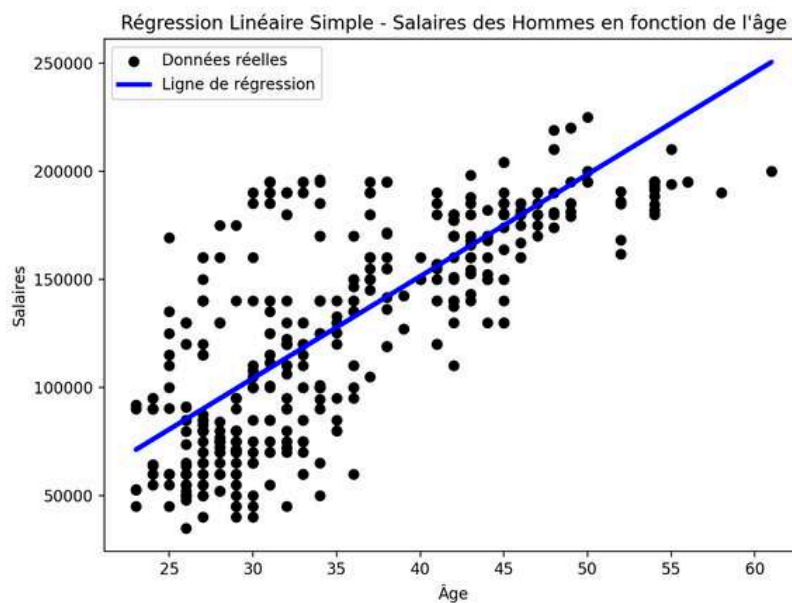
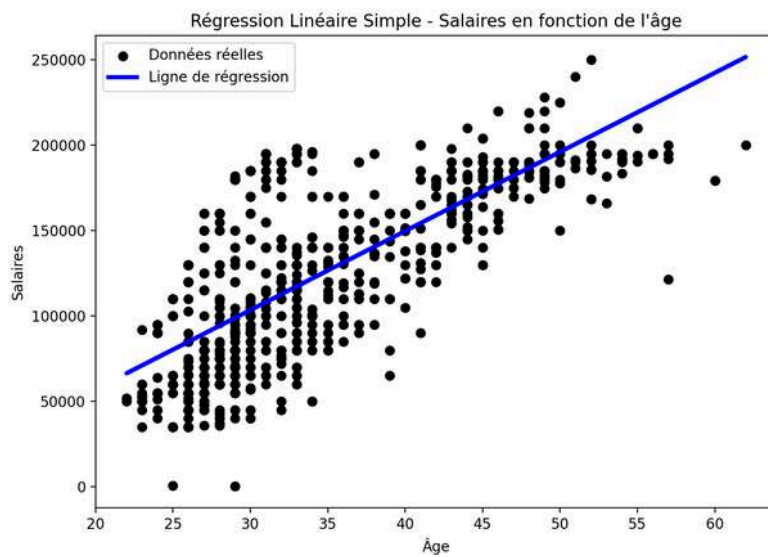


La **rémunération** sera **plus élevée** pour les personnes ayant **plus d'expériences**, homme et femme confondu.

ÉTAPE 2:



Analyse de **l'impact de l'âge** sur les **salaires**.



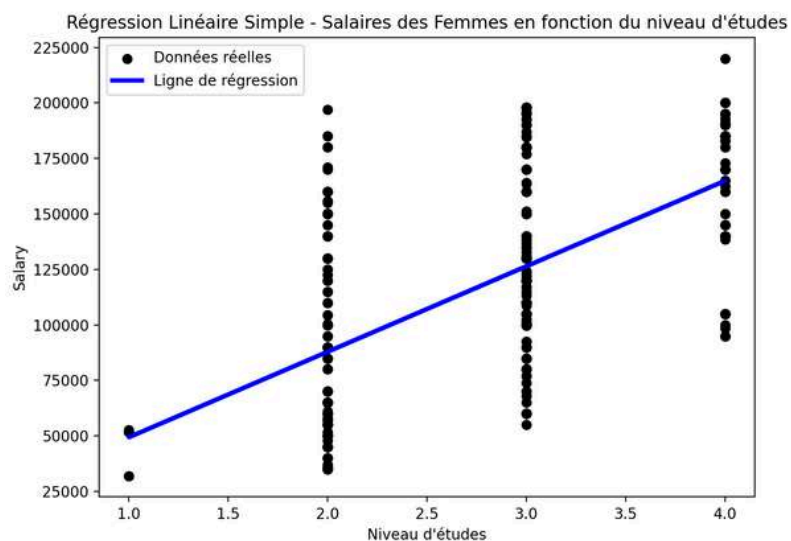
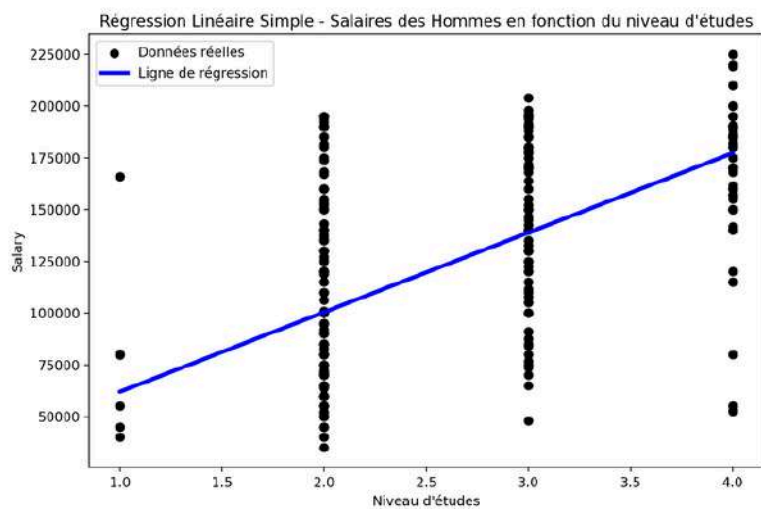
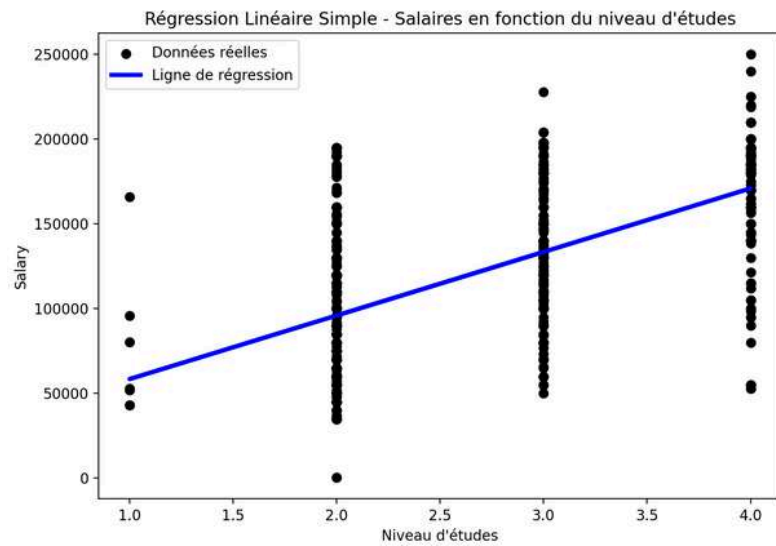
La rémunération sera **plus élevée** pour les personnes étant plus **âgés**, homme et femme confondu.

ÉTAPE 3:



Analyse de **l'impact du niveau d'étude** sur les
salaires.

Pour réaliser cette **régression linéaire**, nous avons eu besoin de transformer les **valeurs discrètes** “High School, Bachelor, Master et PhD”, en **valeurs continues**. Ainsi nous avons attribué 1 à High School, 2 à Bachelor, 3 à Master et 4 à PhD.

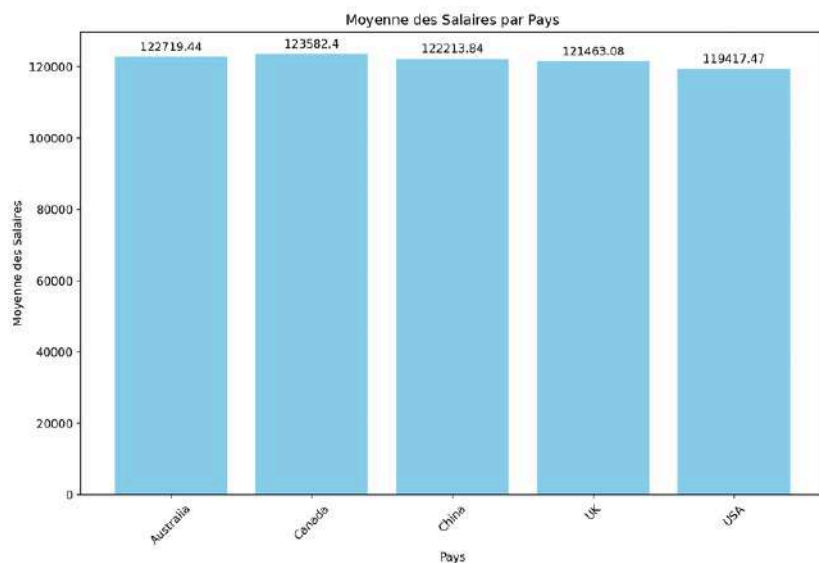


La rémunération sera **plus élevée** pour les personnes ayant fait **plus d'études**, homme et femme confondu.

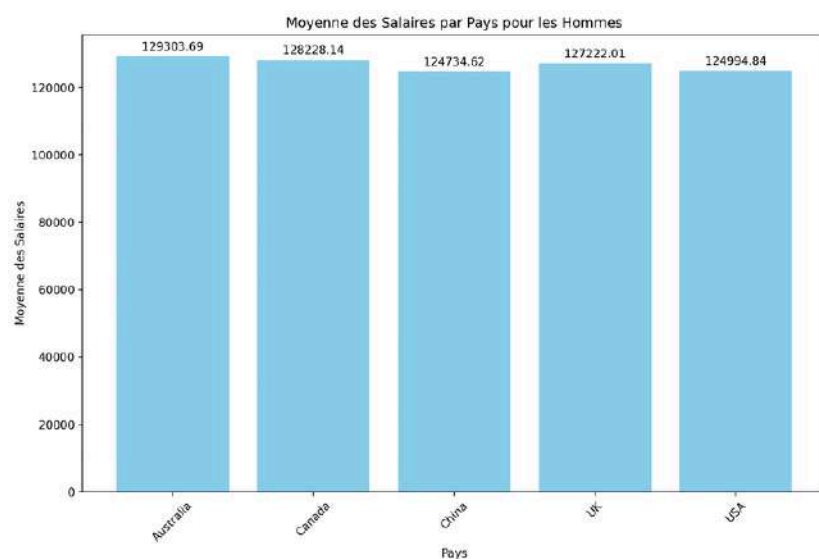
ÉTAPE 4:

A decorative graphic consisting of a large, light gray circle in the center. Overlapping the top-left edge of this circle is a smaller circle with a blue-to-purple gradient. Overlapping the bottom-right edge is another circle of the same size and gradient. The text is centered within the large gray circle.

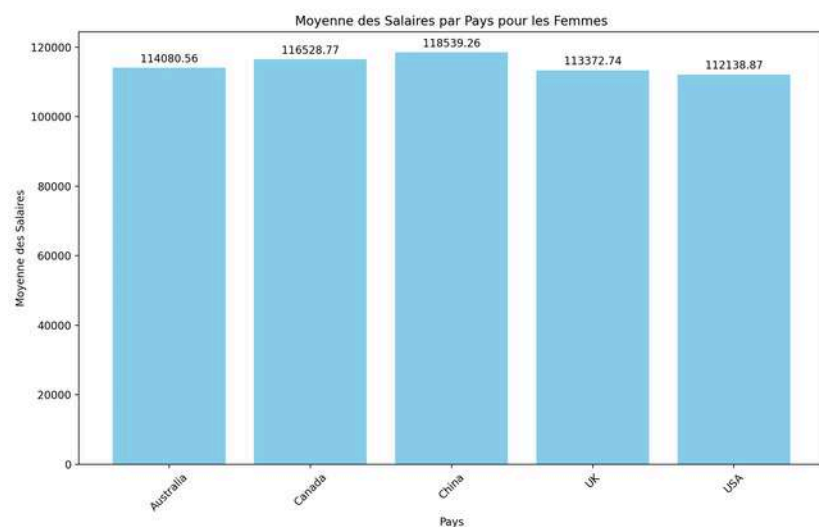
Analyse de **l'impact du pays** sur les **salaires**.



Le pays ayant la moyenne de **salaire la plus élevé**, tout sexe confondu, est le **Canada**, et la **plus basse les USA**.



Le pays ayant la moyenne de **salaire la plus élevé** pour les **hommes** est l'**Australie**, et la **plus basse la Chine**.

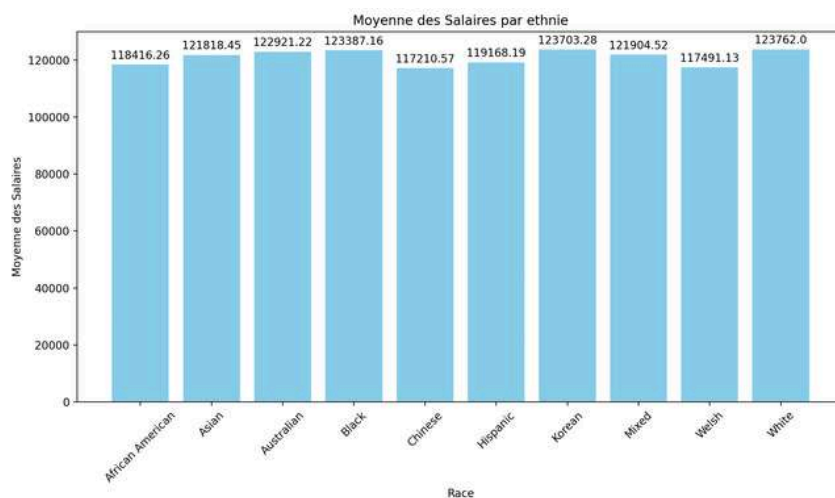


Le pays ayant la moyenne de **salaire la plus élevé** pour les **femmes** est la **Chine**, et la **plus basse les USA**.

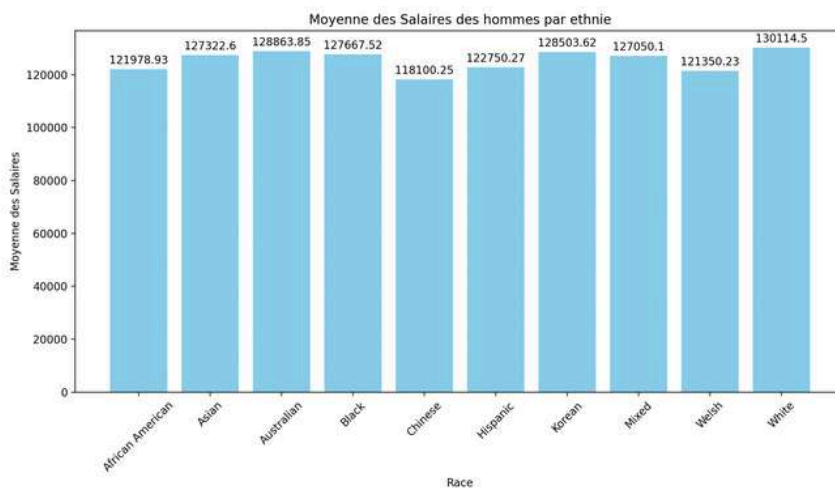
ÉTAPE 5:



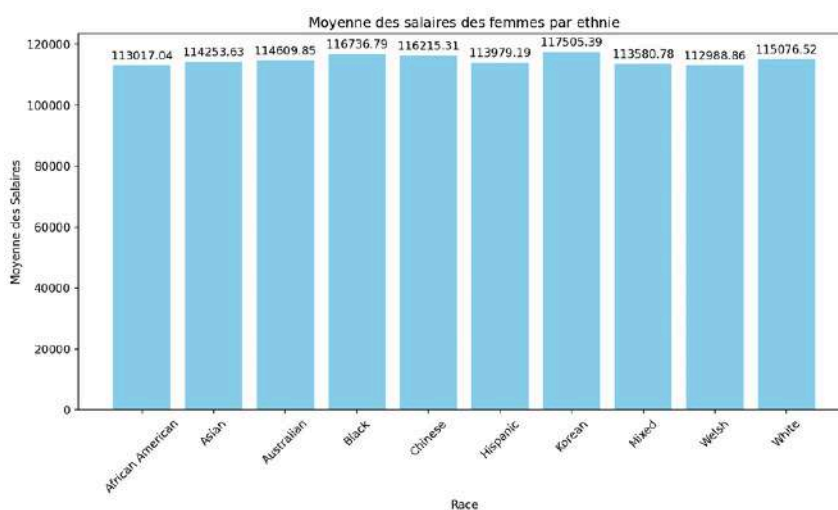
Analyse de **l'impact de l'ethnie** sur les
salaires.



Les personnes, tout sexe confondu, ayant **les salaires les plus élevés** sont d'ethnie **blanche** et **coréenne**. Et les **plus bas**, **chinoise**.



Les hommes ayant **les salaires les plus élevés** sont d'ethnie **blanche** et **australienne**. Et les **plus bas**, **chinoise**.



Les femmes ayant **les salaires les plus élevés** sont d'ethnie **noire** et **coréenne**. Et les **plus bas**, **galloise**.

Interprétation et Conclusion

Discutez de vos trouvailles, en lien avec vos hypothèses initiales, et proposez des recommandations basées sur vos analyses.

Les **tendances du marché du travail** dans les secteurs du **marketing**, de la **data** et de **l'IT** en terme de **rémunération** que nous avons pu prédire sont :

Plus les personnes avanceront dans leurs **carrières professionnelles**, plus elles auront un **salaire élevé**. En effet, nous avons pu voir que **l'âge** et les **années d'expériences** sont des facteurs déterminant dans la rémunération.

Le niveau d'étude est également un facteur déterminant. Afin d'avoir un **salaire plus élevé**, il faut faire des **études supérieures**.

Les pays offrant les **meilleures rémunérations** dans ces métiers sont pour les **hommes l'Australie** et pour les **femmes, la Chine**.

À l'inverse, le pays offrant les **moins bonnes rémunérations** sont pour les **hommes la Chine** et pour les **femmes les USA**.

Et pour finir, **l'ethnie** la **mieux** rémunérée au sein des femmes est **coréenne** et des hommes, **blanche**. À l'inverse, l'ethnie la **moins bien** rémunérée au sein des **femmes** est **galloise** et des **hommes, chinoise**.

MERCI !