

Formez des groupes de 5 personnes et choisissez un des projets suivants. Les groupes peuvent être formés en fonction des affinités entre les étudiants.

Projet 1 : analyse descriptive des données

1. Importation des données sur R sous forme de data frame,
2. Vérification des données,
3. Calculez les statistiques descriptives pour les prix d'ouverture et de fermeture : moyenne, écart type, minimum, maximum, quartiles, etc,
4. Obtenez des informations sur la distribution des prix (ouverture et fermeture),
5. Tracez des graphiques pour visualiser les tendances temporelles des prix d'ouverture et de fermeture sur la période étudiée,
6. Comparez les distributions des prix d'ouverture et de fermeture en utilisant des graphiques ou des tests statistiques,
7. Calculez et examinez la volatilité des prix en utilisant l'écart entre les prix d'ouverture et de fermeture
8. Préparez un rapport décrivant les étapes, les analyses effectuées et les conclusions tirées de l'analyse descriptive des données.

Projet 2: réaliser une ACP sur les données

1. Importation des données sur R sous forme de data frame,
2. Vérification des données,
3. Standardisation des données,
4. Calcul des composantes principales (vous pouvez utiliser la fonction `prcomp()`)
5. Visualisation des résultats (bi-plots),
6. Préparez un rapport décrivant les étapes effectuées et les conclusions tirées de l'ACP.

Projet 3: Analyse de la corrélation entre le prix de fermeture et le prix d'ouverture des données

1. Importation des données sur R sous forme de data frame,
2. Vérification des données,
3. Calcul de la corrélation entre les prix d'ouverture et de fermeture
4. Effectuez des tests statistiques supplémentaires si nécessaire pour évaluer la signification de la corrélation entre les prix d'ouverture et de fermeture, comme le test de corrélation de Pearson,
5. Préparez un rapport décrivant les étapes de l'analyse de la corrélation entre le prix de fermeture et le prix d'ouverture des données.

Projet 4: Prédiction des variations entre prix d'ouverture et de fermeture en développant un modèle de régression.

1. Importation des données sur R sous forme de data frame,
2. Vérification des données,
3. Organisez les données dans un format adapté pour la modélisation, en créant une variable représentant la variation (différence) entre les prix de fermeture et d'ouverture,

4. Division des données en ensembles d'apprentissage et de test (Séparez les données en un ensemble d'apprentissage pour entraîner le modèle de régression et un ensemble de test pour évaluer les performances du modèle),
5. Utilisez des techniques de régression linéaire, ou d'autres modèles de régression disponibles dans R, pour entraîner un modèle capable de prédire les variations entre les prix d'ouverture et de fermeture, (la fonction **lm()**)
6. Utilisez l'ensemble de test pour évaluer les performances du modèle de régression. Mesurez des métriques telles que le coefficient de détermination (R^2), l'erreur quadratique moyenne (RMSE), ou d'autres métriques d'évaluation de la performance du modèle. (la fonction **predict()**)
7. Effectuez une validation croisée pour vérifier la robustesse et la généralisation du modèle.
8. Préparez un rapport décrivant les étapes de la prédiction des variations entre prix d'ouverture et de fermeture en développant un modèle de régression.

Projet 5: analyse de la tendance à long terme avec des techniques de régression linéaire

1. Importation des données sur R sous forme de data frame,
2. Vérification des données,
3. Calculez une variable de tendance à long terme en utilisant des techniques de régression linéaire des prix de fermeture en fonction du temps des données afin d'estimer la pente (la tendance) de la ligne de régression.
4. Tracez les lignes de régression pour les prix de fermeture de chaque action pour visualiser les tendances à long terme.
5. Préparez un rapport décrivant les étapes de l'analyse de la tendance à long terme avec des techniques de régression linéaire.

Projet 6 : Modélisation et prévision avancée des séries temporelles

Analyser les données historiques de prix et développer un modèle pour prédire les prix futurs à l'aide de techniques de séries temporelles.

1. Importation des données sur R sous forme de data frame,
2. Vérification des données,
3. Calculez les statistiques descriptives des prix d'ouverture et de fermeture et analysez leur distribution,
4. Appliquez des transformations comme la différenciation pour rendre la série temporelle stationnaire si nécessaire
5. Utilisez une méthode de décomposition (comme **stl()** dans R) pour séparer la tendance, la saisonnalité et le résidu des prix de fermeture.
6. Utilisez un modèle ARIMA pour modéliser les prix de fermeture.
7. Divisez les données en ensemble d'apprentissage et de test, puis évaluez les performances du modèle sur l'ensemble de test en utilisant des métriques comme l'erreur quadratique moyenne (RMSE),
8. Effectuez des prévisions sur les prix futurs et visualisez-les, en incluant les intervalles de confiance.
9. Préparez un rapport qui décrit les étapes de l'analyse des séries temporelles, les hypothèses posées, le modèle choisi et les prévisions obtenues.

Note : Open c'est les prix d'ouverture et close c'est les prix de fermeture.