

PROJET PERL

PROF : DR ATTA

MEMBRES DU GROUPE :

- BANCE BOUKARE
- ASSI-ADON MARC-IVAN
- BALLO SANATA
- ABDOURAHAMANE HINDOUNE

SOMMAIRE

INTRODUCTION

- I. LES FICHIERS LOGS
- II. DATA VISUALISATION ET L'ANALYSE DE DONNEES
- III. CAS PRATIQUE

CONCLUSION

INTRODUCTION

L'analyse des données est la science qui consiste à examiner les données pour en tirer des informations permettant de prendre des décisions ou d'approfondir les connaissances sur divers sujets. Elle consiste à soumettre les données à des opérations. Ce processus permet d'obtenir des conclusions précises qui nous aident à atteindre nos objectifs, tels que des opérations qui ne peuvent pas être définies au préalable, car la collecte de données peut révéler des difficultés spécifiques.

L'analyse des données est cruciale :

- Une prise de décision éclairée
- Identifier les opportunités et les défis
- Améliorer l'efficacité et la productivité
- Compréhension du client et personnalisation
- Détection de la fraude et sécurité

Il existe plusieurs types d'analyse de données, chacun ayant un objectif et une méthode spécifique. Parlons de quelques types significatifs : **Analyse descriptive, Analyse inférentielle, Analyse prédictive, Analyse normative, etc**

I. LES FICHIERS LOGS

En informatique, un fichier log permet de stocker un historique des événements survenus sur un serveur, un ordinateur ou une application. Ce « journal » présenté sous la forme d'un fichier, ou équivalent, liste et horodate tout ce qui se passe. Les informations contenues permettront ensuite de mieux comprendre les usages et résoudre les erreurs.

Initialement, les fichiers log des serveurs Web comme Apache ou Microsoft IIS étaient l'option par défaut pour enregistrer et traiter une erreur de processus. On a rapidement découvert que le fichier journal d'un serveur Web contenait bien d'autres données importantes : **des Informations sur l'accessibilité et la popularité des sites Internet hébergés sur le serveur** ainsi que les données des utilisateurs comme :

- L'heure et date de consultation
- Le nombre de consultations
- La durée de la session
- L'adresse IP et le nom d'hôte de l'utilisateur
- Les informations sur le client demandeur (en général le navigateur)
- Le moteur de recherche utilisé, dont les requêtes
- Le système d'exploitation utilisé
- Les codes d'erreurs

Une entrée classique d'un fichier log d'un serveur Web se présente comme ci-dessous:

```
183.121.143.32 - - [18/Mar/2003:08:04:22 +0200] "GET /images/logo.jpg
HTTP/1.1" 200 512 "http://www.wikipedia.org/" "Mozilla/5.0 (X11; U; Linux i686;
de-DE;rv:1.7.5)"
```

Dans notre travail :

- Source de fichier log : [Applications \(softwareag.com\)](http://Applications (softwareag.com))

II. DATA VISUALISATION ET ANALYSE DE DONNEES

1- DATA VISUALISATION

La visualisation de données est la représentation des données à travers l'utilisation de graphiques communs, tels que des diagrammes, des représentations, des infographies et même des animations. Ces affichages visuels d'informations communiquent les relations de données complexes ainsi que des informations basées sur les données d'une manière facile à comprendre.

Les tableaux de bord sont des outils de visualisation de données efficaces pour le suivi et la visualisation des données à partir de plusieurs sources de données, offrant une visibilité sur les effets des comportements caractéristiques d'une équipe ou d'un comportement adjacent sur les performances. Les tableaux de bord incluent des techniques de visualisation communes, telles que :

- **Tableaux**
- **Graphiques circulaires et diagrammes à barres empilées**
- **Graphiques linéaires et graphiques de surface**
- **Histogrammes**
- **Nuages de points**
- **Cartes de densité**
- **Cartes d'arbre**

1- ANALYSE DE DONNEE

L'analyse des données consiste à analyser les données pour obtenir des informations sur les clients, les produits ou les ventes. Les entreprises collectent des données non structurées et utilisent l'analyse des données pour les structurer et les interpréter.

Voici les meilleurs outils et logiciels d'analyse de données :

Nom	Visualisation de données	Essai gratuit
 Python	Matplotlib, Seaborn et Bokeh.	Logiciel gratuit
 Zoho	Tableaux et graphiques	15-Day Free Trial
Tableau	Graphiques, graphiques et cartes.	Essai gratuit d'un an
Qlikview	Graphiques, diagrammes et cartes interactifs.	30-Day Free Trial
SAS	Rapports et graphiques	14-Day Free Trial

NB : nous utiliserons le langage perle dans notre cas avec les bibliothèques,
GD::Graph::bars , GD::Graph::pie

III. CAS PRATIQUE

NB : nous utiliserons le langage perle dans notre cas avec les bibliothèques,
GD::Graph::bars , GD::Graph::pie

Source de fichier log : [Applications \(softwareag.com\)](http://Applications (softwareag.com))

Etape 1 : Le code de extraire.pl va permettre d'extraire les données qui nous intéresse dans le fichier log data.txt afin de les mettre dans le fichier data2.txt

Etape 2 : le code dans histogramme.pl et D_circulaire.pl vont se servir des données du fichier data2.txt(respectivement colonne2 et colonne1) pour concevoir respectivement un histogramme et un diagramme circulaire

On peut apercevoir ces graphs dans le dossier graphe.

CONCLUSION

L'analyse des graphes nous a permis de constater au niveau du diagramme circulaire qu'il y a beaucoup plus de code commençant par 200 que de code commençant par 100, 300, 400, 500. Ce qui signifie que l'accessibilité aux sites se fait facilement.

Quand à l'histogramme, on voit une certaine constante au niveau du nombre de demande effectuée sur le serveur par jour.

Références :

- [1] : [Analyse des données : Définition, types et exemples | QuestionPro](#)
- [2] : [Qu'est-ce qu'un fichier log en informatique et à quoi sert-il ? \(hubspot.fr\)](#)
- [3] : [Qu'est-ce que la visualisation des données ? | IBM](#)