

<b>Nom de la SAE</b>	<b>Collecte automatisée de données web</b>		Semestre / Période	S3
Volume horaire consacré par l'étudiant	Avec enseignant	18h00	en autonomie	3h30
Coéquipiers :	Suzanne Martin-Witkowski			
	Théo Duchateau			

Sujet spécifique	Collecte automatisée des données de produits du site internet Action
Objectifs	<ul style="list-style-type: none"> <li>- Compréhension du fonctionnement de l'API</li> <li>- Requêtage des données de produits</li> <li>- Nettoyage des données et conversion en dataframe</li> <li>- Analyse statistiques</li> <li>- Rédaction d'un rapport (Markdown)</li> </ul>
Livrables	Code Python (requêtage des données et analyses statistiques), rapport (Markdown)

**Bilan de la SAE**

*(Reproduire le tableau autant de fois que de compétences mobilisées dans la SAÉ)*

Compétence	Traiter des données à des fins décisionnelles
Apprentissages critiques sollicités	Identifier et résoudre les problèmes d'intégration de sources complémentaires et hétérogènes
	Comprendre la nécessité de tester, corriger et documenter un programme
	Apprécier l'intérêt de briques logicielles existantes et savoir les utiliser
Composantes essentielles à respecter	En intervenant à toutes les étapes du cycle de vie de la donnée (insertion, modification, extraction, suppression)
	En identifiant les librairies et langages dédiés
	En écrivant un programme correctement structuré et documenté, respectant les bonnes pratiques

Compétence	Analyser statistiquement des données
Apprentissages critiques sollicités	Comprendre l'intérêt des analyses multivariées pour synthétiser et résumer l'information portée par plusieurs variables
Composantes essentielles à respecter	En identifiant et en mettant en œuvre les techniques adaptées aux attentes du client ou de l'instance décisionnaire
	En identifiant et en mettant en œuvre les techniques adaptées aux données complexes (données massives, données mal structurées, flux de données...)

Compétence	Valoriser une production dans un contexte professionnel
Apprentissages critiques sollicités	Savoir défendre ses choix d'analyses
	Prendre conscience de la rigueur requise dans ses productions et dans la communication à leur propos
Composantes essentielles à respecter	En interprétant et contextualisant les résultats (citations, vérification des sources, esprit critique)
	En utilisant la forme de restitution adaptée
	En s'exprimant correctement, aussi bien en français que dans une langue étrangère, à l'oral comme à l'écrit

Compétence	Développer un outil décisionnel
Apprentissages critiques sollicités	Percevoir les enjeux de l'automatisation et de l'interopérabilité d'un ensemble de tâches
	Comprendre le cycle de vie d'un projet informatique

Composantes essentielles à respecter	En mettant en œuvre une structuration des données adaptée à leurs caractéristiques (type, volume...)
	En assurant la qualité des données et minimisant les biais liés à l'incertitude et l'imprécision dans les sources
	En réalisant des solutions de visualisation spécifiques aux données métier

### Ma démarche

Savoirs / connaissances	Savoir-faire	Savoir-être
Fonctionnement de l'API  Module json et request	Réalisation d'un code propre et fonctionnel  Requête de l'API avec le module request  Traite des données avec le module json  Analyses statistiques de données issues d'un site web	Organisation en groupe  Rendu de projet répondant à la demande

### Evaluation du résultat

- Ce que je trouve bien réalisé, pourquoi ?

Le code fonctionne et permet de répondre à la problématique du projet.

- Ce que je n'ai pas bien compris ; ce qui serait à améliorer pour une prochaine fois : pourquoi ? comment ?

Nous avons mis un certain temps avant de comprendre comment aborder le sujet.

### Eléments de preuve, ce que je peux montrer

(Choisir des éléments précis à mettre annexe)

- 1) Code Python et dataframe final
- 2) graphiques des analyses statistiques