



**UNIVERSITÉ HASSAN II**  
**FACULTÉ DES SCIENCES AIN CHOCK CASABLANCA**  
**Département de Mathématiques et Informatique**

**Filière Licence professionnelle**  
**Modélisation Statistique Informatique**

---

**Application d'aide à la décision en mesurant la performance des  
nanosatellites sous format d'un tableau de bord dynamique  
dans la société KSF Space**

---

## **MÉMOIRE DE PROJET**

*Réalisé par :*

**IDRISSI MOUNADI DOHA**  
**BENSAHL HAJAR**

*Encadré par :*

**Pr. DEHBI RACHID**

*Jury*

Pr. Z. RADID

Pr. M. SGHIR

Pr. A. CHIBA

Pr. S. MAGHRI

**08 juillet 2023**

# *Dédicaces*

***Ont dédié ce modeste travail :***

***À nos chères familles et proches amis,***

***Nous tenons à vous dédier ces mots empreints d'amour et de gratitude, car vous avez été les piliers inébranlables de notre parcours.***

***À nos parents, frères et sœurs, conjoints et enfants, nous vous remercions du fond du cœur pour votre soutien indéfectible. Vos encouragements, votre compréhension et votre présence constante ont été une source de motivation inestimable. Vous avez été notre source d'inspiration et de réconfort dans les moments de doute, et votre confiance en nos capacités nous a poussés à donner le meilleur de nous-mêmes.***

***À nos proches amis, vous avez été des compagnons de route précieux. Votre soutien inconditionnel, vos encouragements et vos conseils avisés ont été des moteurs essentiels dans notre parcours. Vos sourires, vos mots d'encouragement et vos échanges enrichissants nous ont donné la force et la détermination nécessaires pour surmonter les obstacles.***

# Remerciements

Nous tenons à exprimer nos sincères remerciements dans le cadre de notre rapport de stage réalisé lors de notre formation en Licence Professionnelle de Modélisation Statistique Informatique à l'Université Hassan II à Casablanca.

Tout d'abord, nous souhaitons exprimer notre gratitude envers notre encadrant académique, Monsieur Rachid Dehbi, pour son soutien et son expertise tout au long de notre stage. Ses conseils éclairés et sa disponibilité ont grandement contribué à l'enrichissement de notre expérience. Grâce à sa guidance, nous avons pu approfondir nos connaissances et développer nos compétences dans le domaine de la modélisation statistique et informatique.

Nous tenons également à remercier chaleureusement notre cheffe de filière, Madame Atika Radid, pour son soutien et son implication dans notre parcours universitaire. Sa passion pour l'enseignement et sa volonté de voir ses étudiants réussir ont été une source d'inspiration pour nous.

Nous souhaitons exprimer notre reconnaissance envers les personnes suivantes :

- Monsieur Sghir : Nous le remercions infiniment pour son encadrement attentif, son expertise et son dévouement tout au long de notre stage. Sa bienveillance et ses conseils avisés ont grandement contribué à notre apprentissage.
- Monsieur Chiba Zouhair : Nous tenons à le remercier pour son soutien constant, son partage de connaissances et son engagement à nous aider à relever les défis du stage. Sa patience et son expertise ont été d'une valeur inestimable pour notre progression.
- Monsieur Maghri : Nous lui sommes reconnaissantes pour ses enseignements de qualité, sa disponibilité et son accompagnement tout au long de notre formation. Son expertise et son dévouement ont été très appréciés.

Nous aimerions également exprimer notre gratitude envers Monsieur Kayyali Mohammed, le créateur fondateur de KSF Space et notre encadrant de stage. Sa vision, son soutien et son engagement ont été déterminants dans notre expérience de stage. Nous le remercions sincèrement pour son accueil, sa confiance et les opportunités qu'il nous a offertes.

# Liste des figures

Figure 1 : Logo de KSF Space .....	3
Figure 2 : Organigramme de KSF Space .....	3
Figure 3 : Exemple de certificat NEP.....	5
Figure 4 : Modèle en cascade .....	8
Figure 5 : Modèle de cycle en V.....	9
Figure 6 : Méthode Agile .....	11
Figure 7 : Les perspectives du tableau de bord prospectif .....	11
Figure 8 : les différents types de tableaux de bord .....	12
Figure 9 : Diagramme en cas d'utilisation .....	15
Figure 10 : Diagramme de classe .....	16
Figure 11 : Modèle conceptuel de données.....	17
Figure 12 : Modèle logique de données.....	18
Figure 13 : Logo de power AMC.....	19
Figure 14 : Code de cleaning.....	22
Figure 15 : Code de scraping.....	23
Figure 16 : Logo de visual studio .....	24
Figure 17 : Logo de MySQL Workbench .....	24
Figure 18 : Logo de HTML .....	24
Figure 19 : Logo de CSS .....	25
Figure 20 : Logo de Javascript.....	25
Figure 21 : Logo de Chart.js .....	25
Figure 22 : Logo de bootstrap .....	26
Figure 23 : Logo de python .....	26
Figure 24 : Logo de flask .....	26
Figure 25 : Logo de SQL.....	27
Figure 26 : Authentification .....	27

Figure 27 : Page d'accueil.....	28
Figure 28 : Indicateur1 (DNL).....	29
Figure 29 : Indicateur2 (DM).....	29
Figure 30 : Indicateur3 (DO) .....	30
Figure 31 : Indicateur4 (TNO) .....	31
Figure 32 : Indicateur5 (MJN) .....	31
Figure 33 : Indicateur6 (DN) .....	32
Figure 34 : Indicateur7 (PNN) .....	32
Figure 35 : Indicateur8 (RSN).....	33
Figure 36 : Tableau des variables quantitatives .....	35
Figure 37 : la frequence de la variable country.....	36
Figure 38 : Distribution de nanosatellites par pays .....	36
Figure 39 : Répartition de nanosatellites par niveau.....	37
Figure 40 : Répartition des types de catalogue des nanosatellites .....	37
Figure 41 : Répartition des catégories des nanosatellites par pays .....	38
Figure 42 : Répartition des licences des nanosatellites .....	39
Figure 43 : Table de contingence .....	40
Figure 44 : Table profil-ligne .....	41
Figure 45 : Table profil-colonne.....	43
Figure 46 : Test KHI-DEUX .....	43
Figure 47 : Matrice d'inertie .....	44
Figure 48 : Valeurs propres .....	45
Figure 49 : Coordonnées de la ligne AFC.....	46
Figure 50 : Coordonnées de la colonne AFC.....	47
Figure 52 : Coordonnées de la ligne AFC sans Australie .....	48
Figure 51 : Coordonnées de la colonne AFC sans Australie.....	48

# Les abréviations

PFE	Projet de Fin D'Etudes
HTML	HyperText Markup Language
KSF	Kayyali Space Fondation
NEP	Nanosatellite Engineering Professional
UML	Unifed Modeling Language
CSS	Cascading Style Sheets
PD	Pandas
HTTP	HyperText Transfer Protocol
JSON	Javascripts Object Notation
SQL	Structred Query Language
URL	Uniform Resource Locator
DNL	Distribution des Nanosatellites
DM	Distribution de Masse
DO	Distribution des Organisations
TNO	Types des Nanosatellites par Organisation
MJN	Mise à Jour des Nanosatellites
DN	Distribution des Nations
PNN	Pourcentage des Nanosatellites par Nation
RSN	Répartition de statut des Nanosatellites

# Table des matières

*Dédicaces*

*Remerciements*

*Liste des figures*

*Les abréviations*

<i>Introduction .....</i>	<i>1</i>
<i>Chapitre 1 : Contexte général du projet et problématique.....</i>	<i>2</i>
1 Introduction.....	3
2 Présentation de l'organisation :.....	3
2.1 L'organigramme de KSF Space .....	3
3 Les différents services de l'organisation :.....	4
3.1 Service de conseil en fabrication de petits satellites.....	4
3.2 Service de lancement .....	4
3.3 Services adorables pour les projets de R&D universitaires .....	4
3.4 Programme d'expérimentation KSF .....	4
3.5 Certificat NEP "Nanosatellite Engineering Professional" .....	5
4 Les membres du conseil d'administration .....	5
5 Présentation du projet.....	6
5.1 Définition du projet .....	6
6 Cahier de charges .....	6
6.1 Problématique .....	6
6.2 Besoins fonctionnels.....	6
6.3 Besoins non fonctionnels.....	7
6.4 Les indicateurs de performance adoptés .....	7
7 Cycle de développement du projet .....	8
7.1 Le modèle en cascade .....	8
7.2 Le modèle cycle en V .....	9
7.3 Méthode Agile .....	10
8 Le tableau de bord et ces types :.....	11
8.1 Les types de tableaux de bord :.....	12
8.1.1 Tableau de bord stratégique.....	12
8.1.2 Tableau de bord de contrôle des couts .....	12

8.1.3	Tableau de bord de performance .....	12
<i>Chapitre 2 : Analyse et conception de l'application.....</i>		<i>13</i>
1	Le langage UML .....	14
1.1	Diagramme en cas d'utilisation .....	14
1.2	Diagramme de classe.....	15
2	Méthode Merise.....	16
2.1	Modèle conceptuel des données .....	17
2.2	Modèle logique des données .....	18
3	L'outil adapte pour l'analyse .....	19
<i>Chapitre 3 : Réalisation et mise En œuvre De l'application .....</i>		<i>20</i>
1	Introduction .....	21
2	Nettoyage et collecte des données .....	21
2.1	La source des données .....	21
2.2	Nettoyage des données (cleaning) .....	21
3.1	Web scraping .....	22
3	Les outils de développement utilisés .....	24
3.1	Visual studio .....	24
3.2	My SQL.....	24
3.3	HTML .....	24
3.4	CSS .....	25
3.5	Java script .....	25
3.6	Chart.js.....	25
3.7	Bootstrap.....	26
3.8	Python .....	26
3.9	Flask.....	26
3.10	SQL.....	27
4	Les interfaces graphiques de l'application.....	27
4.1	Authentification.....	27
4.2	Page d'accueil .....	28
4.3	Le premier Indicateur .....	28
4.4	Le deuxième indicateur .....	29
4.5	Le troisième indicateur .....	30
4.6	Le quatrième indicateur .....	30
4.7	Le cinquième indicateur .....	31
4.8	Le seizième indicateur .....	32



4.9	Le septième indicateur .....	32
4.10	Le huitième indicateur.....	33
<i>Chapitre 4 : Analyse statistique des Nanosatellites.....</i>		<i>34</i>
1	Introduction.....	35
2	Etude préliminaire de données .....	35
2.1	Tableau des variables quantitatives.....	35
2.1.1	Variable quantitative : country .....	36
2.1.2	Variable quantitative : level .....	37
2.1.3	Variable quantitative : catalogue.....	37
2.1.4	Variable qualitative : catégorie .....	38
2.1.5	Variable quantitative : license .....	38
3	Analyse factorielle des correspondances .....	39
3.1	Table de contingence.....	39
3.2	Table profil-ligne .....	41
3.3	Table profil-colonne.....	42
3.4	Test de KHI-DEUX.....	43
3.5	Matrice d'inertie.....	44
3.6	Les valeurs propres.....	44
3.7	Cordonnée de la ligne AFC.....	45
3.8	Cordonnée de la colonne AFC .....	47
3.9	Cordonnée de la ligne et colonne AFC sans Australie .....	48
<i>Conclusion.....</i>		<i>49</i>

## INTRODUCTION

Ce rapport représente notre projet de fin d'études réalisé dans le cadre de la licence professionnelle Modélisation-Statistique-Informatique. L'objectif de ce projet était de concevoir et développer une application web interactive d'aide à la décision qui mesure la performance d'un système d'information concernant les nanosatellites présente sous format d'un tableau de bord. Ce tableau de bord a été réalisé au sein de l'organisation KSF Space Fondation à Rabat ont réalisée des tâches de design et d'insertion des images dans différents sites comme Pinterest, Behance, 500 px.

Au cours de ce projet, on va s'intéresser également à l'analyse statistique des données de la base de données de nanosatellites et le processus de conception d'un tableau de bord interactif est organisé en cinq chapitres. Le premier chapitre mettra en avant l'organisation KSF Space et ses services, ainsi qu'il se concentrera sur le cadre général du projet, puis les besoins fonctionnels et non fonctionnels de notre projet. Le deuxième chapitre abordera l'analyse et les méthodes utilisées pour la conception de notre application. Le quatrième chapitre présentera la réalisation de notre projet, les outils utilisés, ainsi que les différentes interfaces de l'application, et le dernier chapitre se focalisera sur une étude préliminaire de données en utilisant les statistiques descriptives l'analyse de données en exploitant la méthode de l'AFC en utilisant la bibliothèque pandas avec python.

Avec l'émergence croissante des nanosatellites et leur utilisation dans divers domaines tels que la communication, l'observation de la terre et la recherche scientifique, il devient essentiel de disposer d'outils permettant de gérer et d'analyser efficacement les données émises par ces nanosatellites. C'est dans ce contexte le projet pour objectif de développer une application web offrant une visualisation claire et interactive des informations liées aux nanosatellites.

# Chapitre 1 : Contexte général du projet et problématique

# 1 Introduction

Ce chapitre, représente la société ou on a effectué notre stage. KSF SPACE, ses principaux services et les membres clés de son administration qui jouent un rôle essentiel dans son bon fonctionnement. Par la suite, il existe une exploration du cadre général projet réalise, en détaillant le cahier des charges de notre application. Nous examinons également les différents cycles de développement du projet, ainsi que les divers types de tableaux de bord.

## 2 Présentation de l'organisation :

KSF Space Fondation est une organisation américaine à but non lucratif enregistrée dans le Delaware #560531 offres un soutien dans la réalisation ou sont en train de réaliser de petites missions satellites, avec ceux qui ont une expérience limitée à une fraction du frais habituels. Cela permet aux institutions du monde entier d'explorer et d'utiliser tout ce que le marché a à offrir, sans sacrifier une part importante de leur budget.



Figure 1 : Logo de KSF Space

### 2.1 L'organigramme de KSF Space

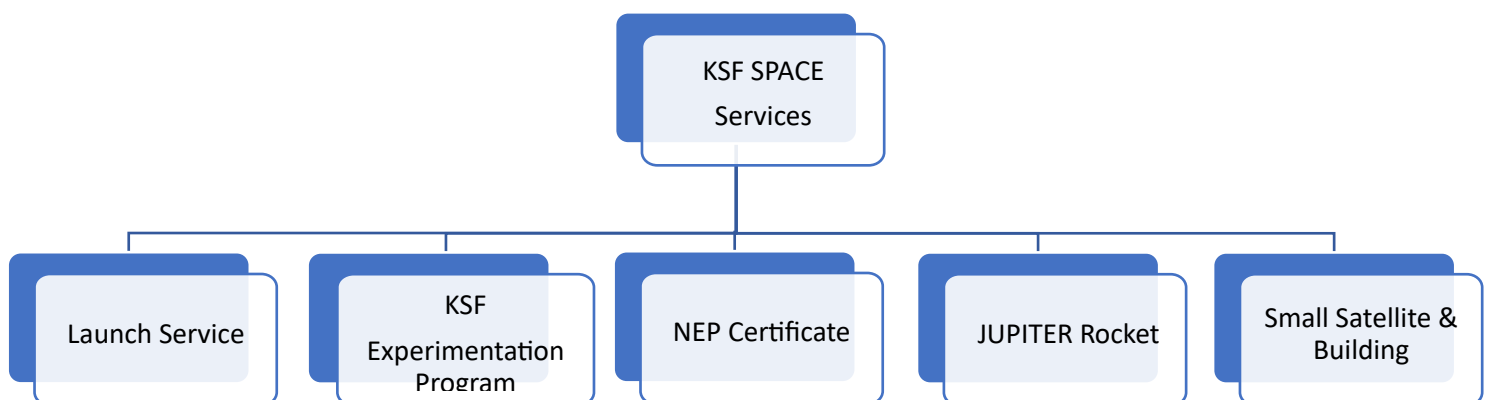


Figure 2 : Organigramme de KSF Space

## **3 Les différents services de l'organisation :**

### **3.1 Service de conseil en fabrication de petits satellites**

Le comité technique international de la KSF Space Fondation (KSF), composé d'ingénieurs professionnels expérimentés, de chercheurs, de scientifiques et d'experts en la matière, peut vous aider et travailler avec vous lors de votre prochaine mission spatiale.

KSF a signé de nombreux accords avec des acteurs clés de l'industrie et de l'enseignement dans le domaine des nanosatellites. KSF peut contribuer à votre prochaine mission spatiale en faisant correspondre les services les plus appropriés du marché aux besoins de votre mission.

### **3.2 Service de lancement**

À la KSF Space Fondation, nous pouvons trouver et gérer des opportunités de partage de trajet avec un réseau de fournisseurs de lancements internationaux. Nous pouvons aider les institutions à lancer leur propre expérience, à des prix inférieurs à ceux du marché.

### **3.3 Services adorables pour les projets de R&D universitaires**

- Construction, intégration, environnement d'essai en apesanteur et possibilités de lancement de nanosatellites "CubeSat"
- Soutenir vos activités universitaires en construisant votre propre station au sol pour la R&D et recevoir des signaux et des images en direct des principaux satellites météorologiques et climatiques en orbite
- Soutien et formation à la construction d'une station au sol pour suivre les nanosatellites ou les cubesats en orbite LEO Nous envoyons vos expériences scientifiques biologiques dans l'espace à bord de l'ISS
- Diffusion des chances avec l'équipage de la Station spatiale internationale (ISS) en orbite pour les étudiants des universités et des collèges du monde entier "par l'intermédiaire du partenaire de KSF Space"
- Cours de formation disponibles sur le GPS, le GS, la construction de NanoSat...etc

### **3.4 Programme d'expérimentation KSF**

Certaines institutions ne peuvent pas facilement assumer les coûts de développement, de fabrication et de lancement de leur propre mission spatiale, en particulier dans les pays en développement.

KSF, a développé un programme spécial intitulé « Experimentation Opportunity Program ».

Ce programme est dédié à permettre aux institutions des pays en développement de réaliser leurs propres aspirations en matière de missions spatiales. La portée du programme permet aux institutions des pays en développement de collaborer et de participer avec d'autres institutions qui travaillent actuellement et développent leurs propres missions.

### 3.5 Certificat NEP “Nanosatellite Engineering Professional”

La filière de certificat NEP s'adressera aux ingénieurs et experts de l'aérospatiale et sera reconnue par les principales entreprises, organisations, fondations et agences spatiales.

Un certificat NEP sera reconnu et accepté au niveau mondial et contribuera à garantir de meilleures opportunités d'emploi dans l'industrie des satellites à l'avenir. De nombreux emplois futurs nécessiteront obligatoirement un certificat NEP - si certains ingénieurs en satellites, télécommunications ou électronique ne sont pas certifiés NEP, ils risquent de ne pas être qualifiés pour le poste, quel que soit le nombre d'années d'expérience qu'ils possèdent.



Figure 3 : Exemple de certificat NEP

## 4 Les membres du conseil d'administration

La Fondation spatiale KSF est dirigée par des membres professionnels et un conseil d'administration ayant des réalisations notables dans les domaines de la science, de l'ingénierie, de l'espace et de l'aviation. Le conseil d'administration et les membres des comités sont tous bénévoles.

Les premiers responsables du conseil d'administration, ainsi que le président supervisent et gèrent quatre comités :

- Le comité de la conférence MNSAT
- Comité de l'éducation
- Comité R&D
- Comité JUPITER

Les membres du conseil d'administration sont :

- Kayyali Mohamed, Président, USA
- Sunny Narayanan, responsable de la biologie Université de Floride, États-Unis
- Erwin te Beek, responsable de l'ingénierie, Europe
- Mohamed Essaidi, responsable de la conférence MNSAT, Maroc
- Jacob Wade, responsable NEP, DoD USA
- Jayakumar Venkatesan, directeur de la technologie, Inde
- Brian Kaplinger, responsable du comité JUPITER
- Université du Kansas, département aérospatial.

## 5 Présentation du projet

### 5.1 Définition du projet

Le projet est "la conception une application d'aide à la décision sous forme d'un tableau de bord dynamique de nanosatellites" a été réalisé au sein de l'organisation KSF Space. Cette application a pour objectif de fournir une application web permettant d'analyser, visualiser et prendre des décisions basées sur les données relatives aux nanosatellites.

## 6 Cahier de charges

### 6.1 Problématique

La problématique centrale de notre projet se représente dans l'absence de la digitalisation et l'optimisation de notre base de données, pour résoudre ce problème on a réalisé une application qui permettra d'optimiser l'utilisation des données par filtration selon le besoin afin de garantir une exploitation plus efficace et pertinente des informations contenues dans la base de données en évitant les risques de plantage de serveur.

### 6.2 Besoins fonctionnels

- Importer les données des nanosatellites à partir de sources externes.
- Mettre à jour les données existantes en fonction des nouvelles informations.
- Supprimer les données obsolètes ou non pertinentes.
- Offrir la possibilité de personnaliser l'affichage des indicateurs en fonction des besoins spécifiques.
- Afficher de manière dynamique les indicateurs relatifs à l'organisation et à la distribution des types de nanosatellites.
- Permettre la visualisation des mises à jour des informations sur les nanosatellites.
- Fournir des fonctionnalités de filtrage pour sélectionner les données pertinentes en fonction des besoins.
- Calculer des statistiques et des métriques basées sur les données des nanosatellites.
- Utiliser les indicateurs et les analyses pour évaluer la performance et prendre des décisions éclairées.
- Facilitation du process de gestion.

### 6.3 Besoins non fonctionnels

Ce sont des besoins pour identifier les exigences suivantes pour notre application :

- Garantir des temps de réponse rapides lors de l'affichage des données et des calculs statistiques.
- Gérer efficacement de grandes quantités de données liées aux nanosatellites.
- Fournir une interface utilisateur intuitive et conviviale pour faciliter la navigation et l'utilisation de l'application.
- La conformité aux réglementations et aux normes de protection des données doit être respectée.

### 6.4 Les indicateurs de performance adoptés

- Distribution de l'année de lancement :  
Cet indicateur représente la distribution des missions en fonction de leur année de lancement, filtrée par la date de lancement.
- Distribution de la masse des nanosatellites :  
Cet indicateur calcule la somme de chaque type de masse entre deux dates de lancement, il donne un aperçu de la taille ou du poids typique des nanosatellites.
- Répartition des organisations :  
Donne un aperçu de la répartition des organisations impliquées dans les missions. Il montre quelles organisations ont mené le plus de missions.
- Les types de catégories de nanosatellites par organisation :  
Cet indicateur mesure la somme des types de catégories en utilisant le nom de l'organisation.
- La mise à jour des nanosatellites :  
Calcule le nombre des nanosatellites mises à jour entre deux dates de lancement différentes.
- Répartition des nations :  
Affiche la répartition des nations impliquées dans les missions. Il permet d'identifier les pays qui participent le plus activement aux missions spatiales.
- Niveau des nanosatellites par organisation :  
Répartition des différents niveaux de nanosatellites (par exemple : national, régional...)
- Répartition des statuts par nation :  
Répartition des missions de nanosatellites selon leur état ou status (par exemple : opérationnel, réentrée, pas de signal) et affiche le pourcentage des différents statuts dans chaque pays.
- Nombre total de missions :  
Cet indicateur représente le nombre total de missions dans l'ensemble de données. Il donne un aperçu du nombre de missions effectuées entre deux dates de lancement différentes.



## 7 Cycle de développement du projet

Le cycle de vie d'une application est constitué de toutes les étapes qui marquent son parcours, depuis l'idée de départ jusqu'à l'abandon du logiciel par ses utilisateurs. Il existe différents types de cycles de développement entrant dans la réalisation d'un logiciel : [le modèle en cascade](#), [Le modèle cycle en V](#) et [La méthode Agile](#). Dans ce projet on a adapté la dernière méthode.

### 7.1 Le modèle en cascade

Le modèle en cascade (en anglais : waterfall model) est un modèle de gestion linéaire qui divise les processus de développement en phases de projet successives. Contrairement aux modèles itératifs, chaque phase est effectuée une seule fois. Les sorties de chaque phase antérieure sont intégrées comme entrées de la phase suivante. Le modèle en cascade est principalement utilisé dans le développement de logiciels. En pratique, plusieurs versions du modèle en cascade sont utilisées. Les modèles les plus courants divisent les processus de développement en cinq phases.

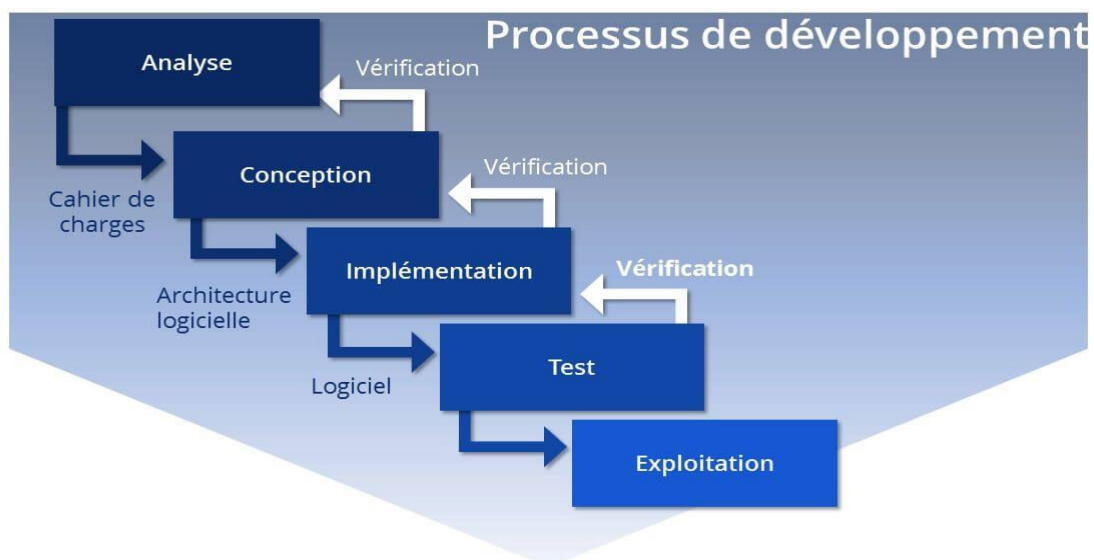


Figure 4 : Modèle en cascade

- **Analyse** : planification, analyse et spécification des besoins.
- **Conception** : conception et spécification du système.
- **Implémentation** : programmation et tests des modules.
- **Test** : intégration du système, tests du système et de l'intégration.
- **Exploitation** : livraison, maintenance, amélioration.

## 7.2 Le modèle cycle en V

Le modèle du cycle en V a été imaginé pour pallier le problème de réactivité du modèle en cascade. Ce modèle est une amélioration du modèle en cascade qui permet en cas d'anomalie, de limiter un retour aux étapes précédentes. Les phases de la partie montante doivent renvoyer de l'information sur les phases en vis-à-vis lorsque des défauts sont détectés afin d'améliorer le logiciel. Dans un premier temps, le cycle en V définit le déroulement d'un projet en phases distinctes qui sont tour à tour détaillées :

- En début de projet, le modèle prévoit une analyse de l'ensemble des besoins relatifs au système envisagé.
- Le projet est ensuite enrichi par l'expression des besoins fonctionnels et non fonctionnels liés à l'architecture du système.
- Puis on passe à la phase de conception du système, lors de laquelle les composants et les interfaces du système sont planifiés.
- Une fois ces étapes franchies, on peut passer à la conception de l'architecture logicielle en détail.

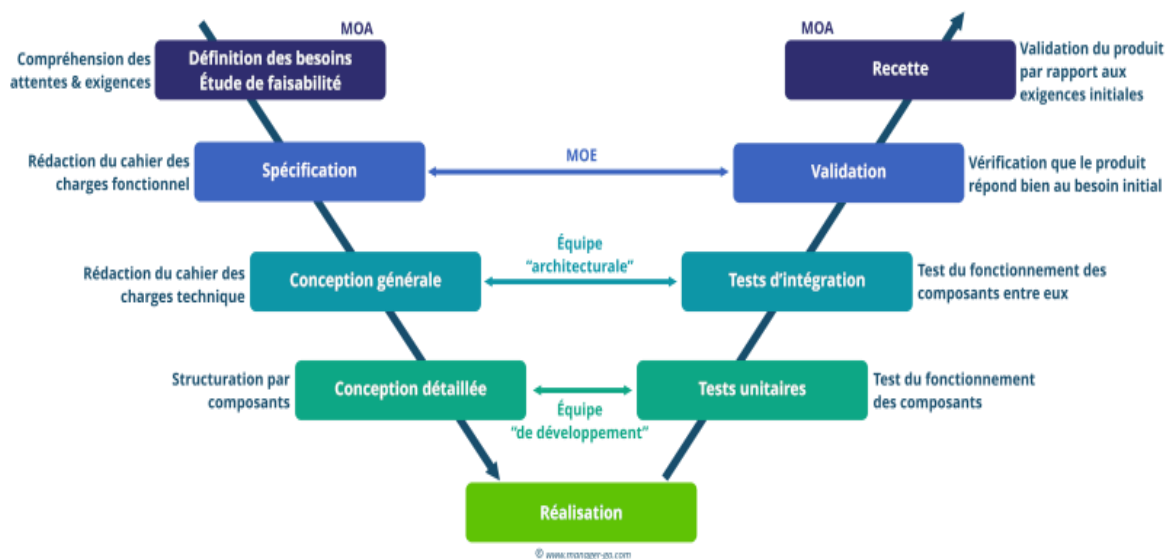


Figure 5 : Modèle de cycle en V

On entre alors dans la phase effective de développement du logiciel en fonction de ce qui a été planifié. Ensuite ce sont les phases d'assurance qualité, qui se réfèrent toujours aux étapes du développement. Le modèle prévoit les tâches suivantes :

- Tests unitaires
- Tests d'intégration
- Intégration système
- La « recette » (ou test d'acceptation)

### 7.3 Méthode Agile

Est une méthode de gestion de projets et de développement qui est très utilisée au sein des agences de marketing et des entreprises technologiques. La spécificité de la méthode agile est son approche par itération. Ce sont des cycles de développement courts, très ciblés, impliquant le client et favorisant la collaboration entre des équipes pluridisciplinaires. L'objectif de la méthode agile est de développer des produits plus rapidement, à moindre coût, et avec un taux de réussite et de satisfaction plus important.

Il existe différentes méthodes Agile et autant de façons de les mettre en œuvre dans une entreprise. Pour autant, une méthodologie commune peut être dégagée. En résumé, développer un produit selon la méthode Agile revient à diviser le travail en une suite d'objectifs atteignables.

Le processus se compose de différentes étapes :

- Le responsable de projet identifie et définit la nature du projet confié à l'entreprise.
- Le responsable et l'équipe fixent les exigences liées au projet. On les appelle User Stories (US) pour Histoire utilisateur. Il s'agit de bien les définir et d'établir leur niveau de priorité ainsi que les personnes les plus à même de les mener à bien.
- Le responsable et l'équipe planifient les itérations, appelées aussi Sprint, c'est-à-dire une suite de cycles courts de développement. A l'intérieur de chaque Sprint sont hiérarchisées les exigences et retours client ainsi que les tâches à accomplir.
- Tous les membres de l'équipe participent à des réunions quotidiennes, appelées aussi « mêlées » ; les échanges durent une quinzaine de minutes et permettent de faire le point sur ce qui a été fait hier et doit être fait aujourd'hui. C'est aussi le moment d'aborder les éventuels points de blocage, les retards, les problèmes... et les solutions.
- A la fin de chaque itération a lieu une revue de Sprint : elle sert à évaluer et livrer au client la version du produit à cet instant « T ». Il est alors temps de récolter le feedback du client, c'est-à-dire ses retours, positifs comme négatifs, et de les traiter de façon réactive.
- Une rétrospective de sprint permet enfin de revenir sur ce qui a été fait, d'en dégager les points forts et les points faibles, de se remettre en question et de perfectionner la méthodologie de travail pour le sprint suivant.



Figure 6 : Méthode Agile

## 8 Le tableau de bord et ces types :

Le balanced scorecard (ou tableau de bord prospectif en français) est un outil managérial permettant de mesurer l'activité d'une entreprise. Il a été développé en 1992 par deux économistes américains Robert S. Kaplan et David Norton. Le balanced scorecard est utilisé par les managers afin de s'assurer que l'action de leur entreprise est conforme aux objectifs fixés sur le long terme. Plutôt que de se focaliser uniquement sur des enjeux financiers, il permet de gérer la performance globale d'une société en s'appuyant sur quatre grandes Perspectives :

- La perspective financière : analyse des bénéfices financiers et des pertes.
- La perspective client : analyse de l'impact de l'entreprise auprès de la clientèle.
- La perspective processus interne : analyse de l'efficacité des processus internes à l'entreprise.
- La perspective apprentissage et développement : analyse de la politique RH de l'entreprise.

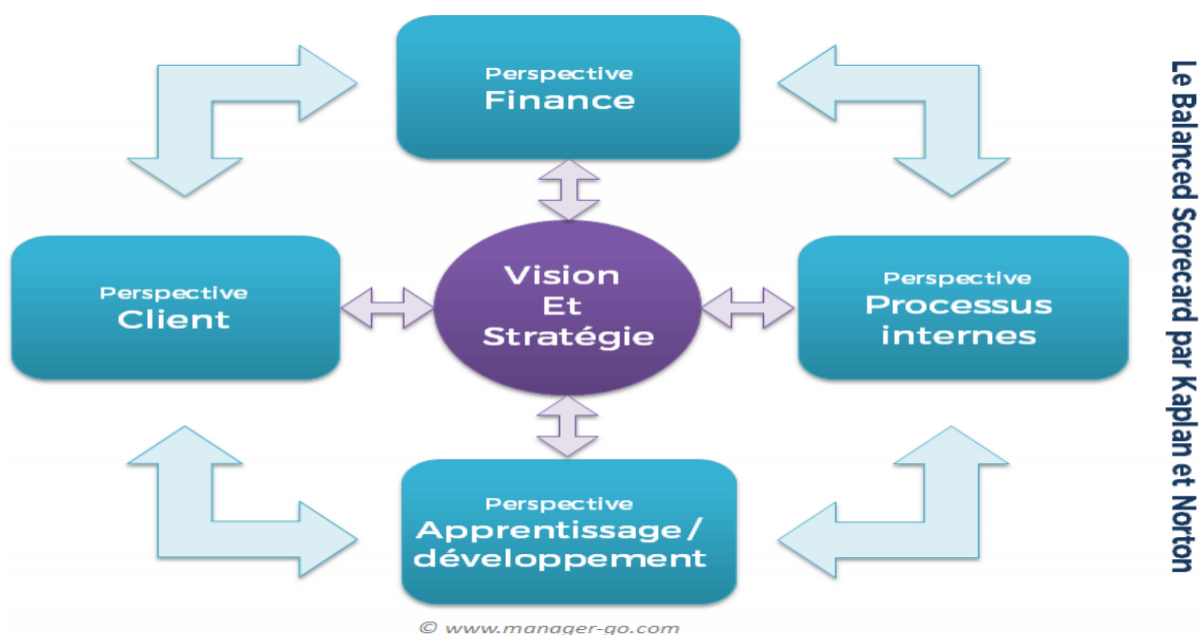


Figure 7 : Les perspectives du tableau de bord prospectif

## 8.1 Les types de tableaux de bord :

Le choix d'un tableau de bord dépend en premier lieu de vos objectifs généraux. Il doit permettre de répondre aux objectifs en présentant les informations qui sont pertinentes. À l'opposé, un tableau de bord mal conçu pourrait vous induire en erreur avec des informations inadéquates, ou vous faire passer à côté de celles qui sont importantes. Il est possible de distinguer 3 types de tableau de bord :

### 8.1.1 Tableau de bord stratégique

Ce type de tableau est conçu pour permettre aux cadres d'avoir une vue globale sur le fonctionnement général de l'entreprise. Le tableau de bord stratégique servira donc surtout aux dirigeants qui ont besoin d'avoir une vue d'ensemble rapide de l'état des lieux, leur permettant ainsi de suivre les mesures de performance par rapport aux objectifs stratégiques de l'entreprise.

### 8.1.2 Tableau de bord de contrôle des coûts

Le principal objectif de ce type de tableau de bord est de piloter le niveau de performance de la gestion de l'entreprise. Il servira donc à évaluer les performances financières pour les comparer aux prévisions et objectifs qui ont été fixés en amont. Avec un tel tableau de bord, vous pourrez identifier les différences entre vos projections (vos objectifs) et la réalité. De plus, ils mettent en avant les données historiques afin d'identifier les tendances, pour ensuite prédire les résultats et fixer les objectifs à atteindre.

### 8.1.3 Tableau de bord de performance

Véritable outil de suivi des performances opérationnelles, ce type de tableau de bord est utilisé dans le but de suivre la progression vers l'atteinte des différents objectifs opérationnels de l'entreprise. En plus du suivi, cet outil vous permet d'identifier les processus internes qui fonctionnent de manière optimale, ainsi que ceux qui sont à améliorer.

	Tableau de bord stratégique	Tableau de bord de contrôle des coûts	Tableau de bord des performances
Objectif	Suivi des mesures de performance par rapport aux objectifs stratégiques de l'entreprise	Gestion de la performance financière de l'entreprise	Suivi et évaluation de la performance des processus opérationnels de l'entreprise
Audience	Les cadres supérieurs	Les cadres et l'équipe administrative	L'équipe responsable de la production
Vision	Moyen et long terme	Moyen terme	Court terme

Figure 8 : les différents types de tableaux de bord

## Chapitre 2 : Analyse et conception de l'application

# 1 Le langage UML

Le langage de modélisation unifié, est un langage de modélisation graphique à base de pictogrammes conçus comme une méthode normalisée de visualisation dans les domaines du développement de logiciel et en conception orientée objet.

## 1.1 Diagramme en cas d'utilisation

Un diagramme de cas d'utilisation est un outil de modélisation graphique utilisé en ingénierie logicielle pour représenter les interactions entre les acteurs (utilisateurs ou systèmes externes) et un système. Il permet de capturer les différents scénarios ou cas d'utilisation auxquels un système peut être confronté.

Le diagramme de cas d'utilisation se compose de plusieurs éléments clés :

**Acteurs** : Les acteurs sont des entités externes qui interagissent avec le système. Dans notre application l'utilisateur est le directeur de KSF Space qui interagit directement avec le système étudié.

**Cas d'utilisation** : représentent des fonctionnalités ou des actions spécifiques que le système peut effectuer en réponse aux demandes des acteurs. Chaque cas d'utilisation est représenté par une ellipse sur le diagramme.

**Relations entre acteurs et cas d'utilisation** : Des lignes sont utilisées pour représenter les relations entre les acteurs et les cas d'utilisation. Ces relations indiquent les interactions possibles entre les acteurs et le système.

.

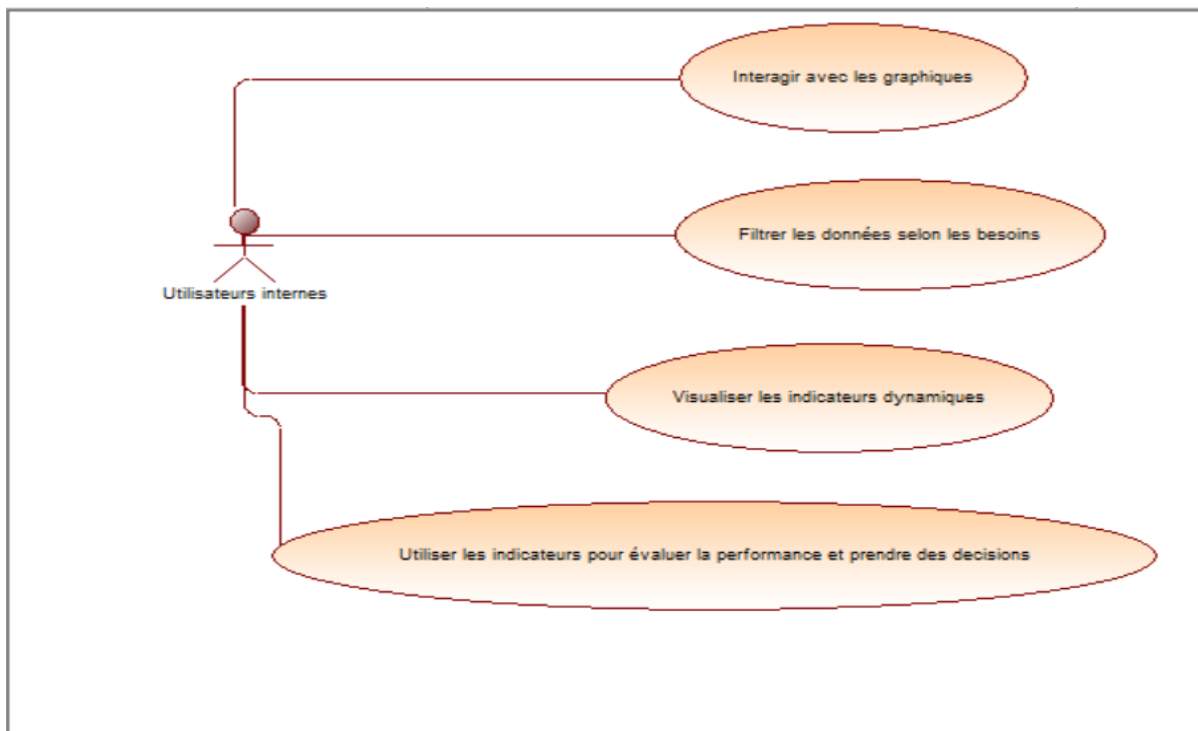


Figure 9 : Diagramme en cas d'utilisation

## 1.2 Diagramme de classe

Un diagramme de classes est un outil de modélisation graphique utilisé en ingénierie logicielle pour représenter les classes, les attributs, les opérations et les relations entre les objets d'un système logiciel. Il offre une vue statique de la structure et des relations des différentes entités du système.

- **Classe** : Une classe est une abstraction qui représente un concept, un objet ou une entité du système. Elle définit les propriétés (attributs) et les comportements (opérations) associés à cette entité. Une classe est généralement représentée sous la forme d'un rectangle divisé en trois sections le nom de la classe, les attributs et les opérations.
- **Attribut** : Un attribut est une caractéristique ou une propriété d'une classe. Il décrit l'état ou les données qu'un objet de cette classe peut stocker. Les attributs sont représentés sous la forme d'une liste dans la section des attributs de la classe, comprenant généralement le nom de l'attribut et son type de données.
- **Opération** : Une opération est un comportement ou une action associée à une classe. Elle représente une méthode ou une fonctionnalité que les objets de cette classe peuvent exécuter. Les opérations sont généralement représentées sous la forme d'une liste dans la section des opérations de la classe, comprenant le nom de l'opération, les paramètres et le type de retour.
- **Relation** : Les relations entre les classes sont utilisées pour représenter les liens et les dépendances entre les objets du système.



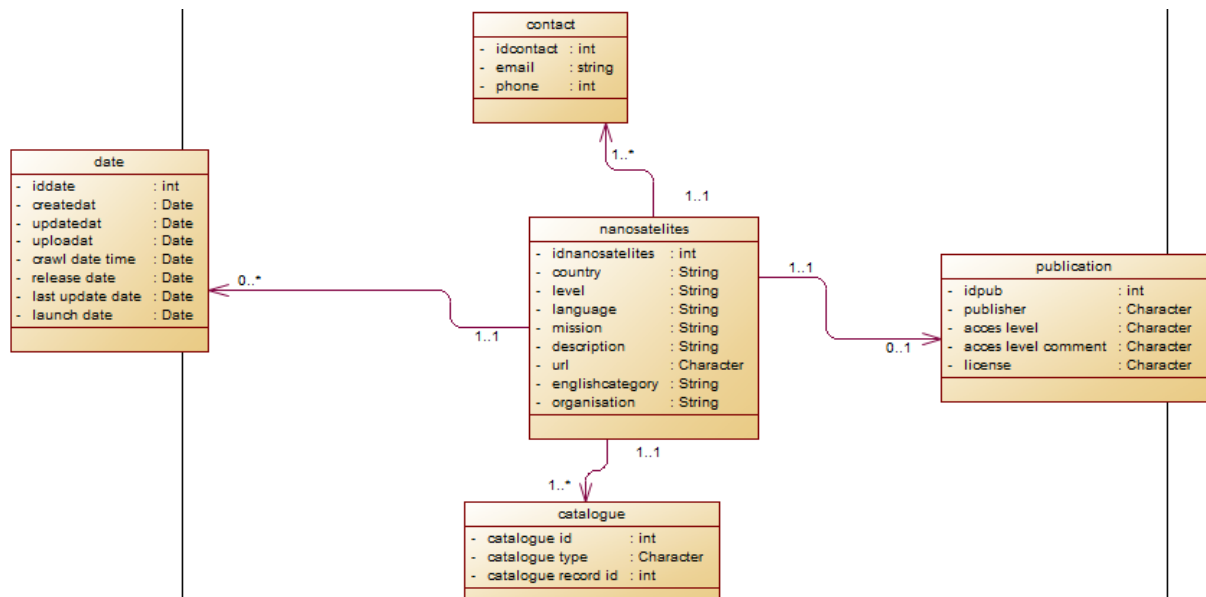


Figure 10 : Diagramme de classe

## 2 Méthode Merise

La méthode Merise est une méthode de conception de systèmes d'information basée sur une approche régulière et progressive. Elle a été développée dans les années 1970 par René Colletti et Hubert Tardieu. La méthode Merise vise à modéliser les données, les traitements et les interactions d'un système d'information en utilisant des diagrammes et des notations graphiques pour concevoir des tables avec des relations pour créer une base de données relationnelle.

Elle possède plusieurs modèles qui sont repartis sur trois niveaux : le niveau conceptuel, le niveau logique et le niveau physique.

## 2.1 Modèle conceptuel des données

Le MCD, ou Modèle Conceptuel de Données, est une représentation visuelle avancée qui vise à formaliser de manière précise les données qui seront utilisées par un système d'information. Il s'agit essentiellement d'une illustration claire des données, permettant de décrire le système d'information à l'aide d'entités compréhensibles. Grâce à un MCD, il est possible de valider et de préciser les règles qui s'appliqueront à la future base de données, en se référant à l'image du MCD suivant.

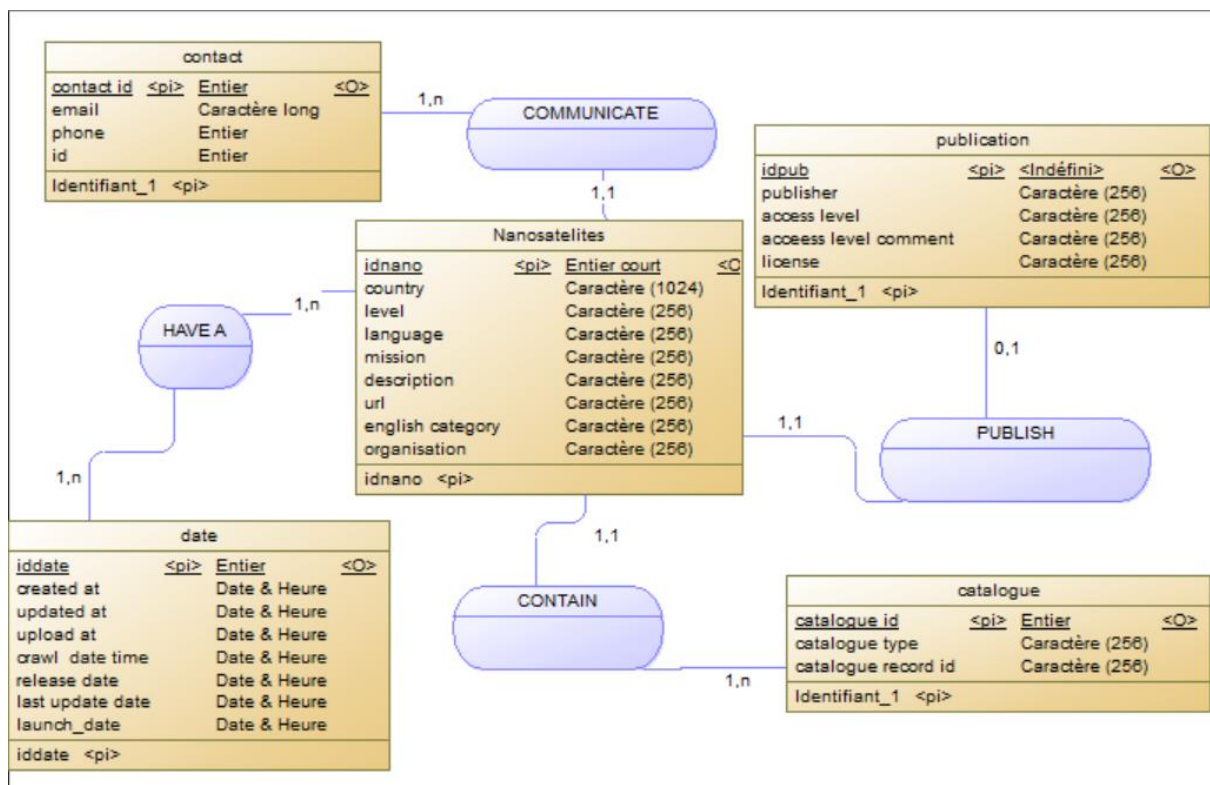


Figure 11 : Modèle conceptuel de données

## 2.2 Modèle logique des données

Le modèle logique des données est une représentation des données d'un système d'information, indépendamment de toute considération technique ou physique. Il définit les entités, les relations et les contraintes qui régissent les données au sein du système. Le modèle logique de données se penche d'avantage sur les détails spécifiques, tels que les types de données, les clés primaires et étrangères, les contraintes d'intégrité, etc. le modèle logique des données permet de traduire le modèle conceptuel en une représentation plus précise et technique, prête à être implémentée dans une base de données qu'on peut l'utiliser lors du développement de l'application.

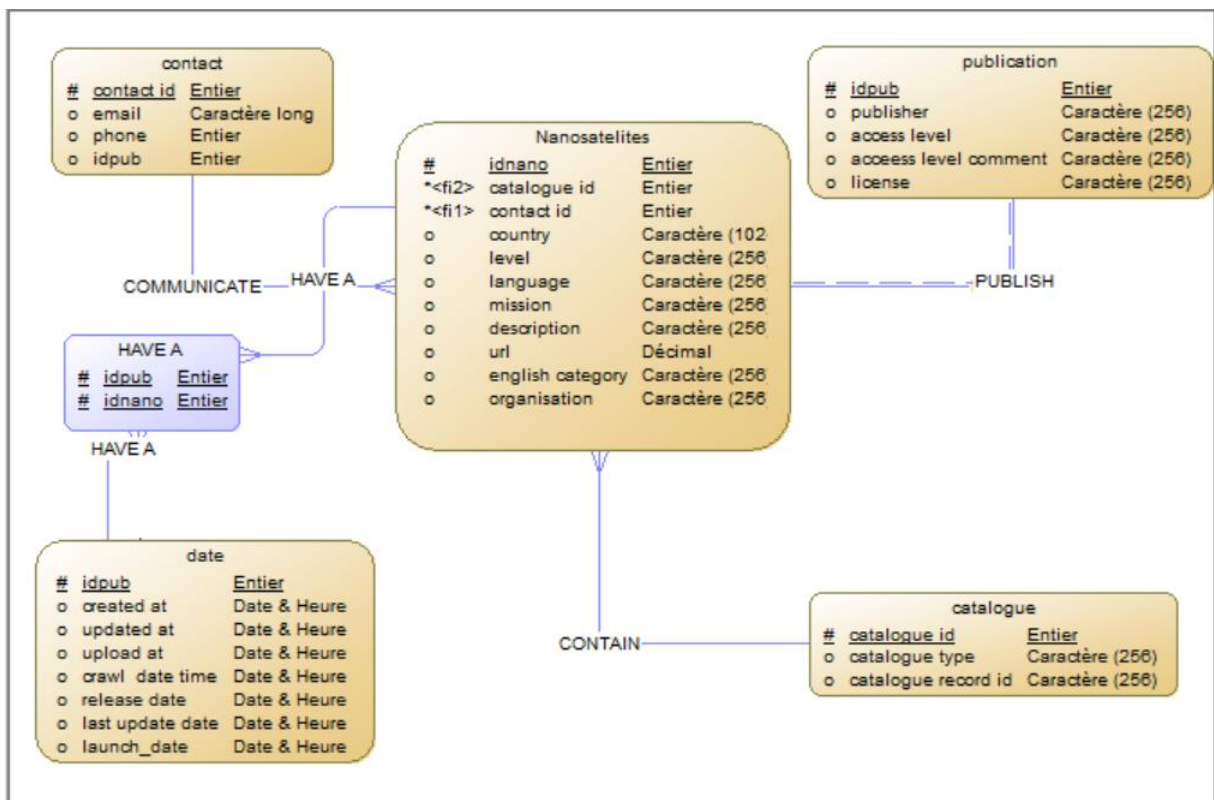


Figure 12 : Modèle logique de données

### 3 L'outil adapte pour l'analyse

Power AMC est un outil de modélisation et de conception de bases de données, largement utilisé dans le domaine de l'ingénierie logicielle. Il offre une interface conviviale permettant de créer et de gérer des modèles conceptuels, logiques et physiques de données. Power AMC propose une gamme complète de fonctionnalités, notamment la création de diagrammes entité-association (ER), la génération automatique de scripts SQL, la rétro-ingénierie à partir de bases de données existantes, la gestion des contraintes et des relations, ainsi que la collaboration en équipe pour le développement de bases de données. Cet outil est largement utilisé par les professionnels de la gestion de données et les architectes de logiciels pour concevoir et maintenir des bases de données de manière efficace et professionnelle.



Figure 13 : Logo de power AMC

## Chapitre 3 : Réalisation et mise En œuvre De l'application

# 1 Introduction

Le chapitre actuel représente une étape essentielle de notre projet, car il se focalise sur la réalisation de notre application. Après avoir mené une étude préliminaire exhaustive et une étude conceptuelle détaillée, nous sommes maintenant prêts à passer à la phase de réalisation concrète de notre solution. Dans ce chapitre, nous présenterons le nettoyage des données et le web scraping et les langages de programmation utilisés et nous inclurons des captures d'écran des principales interfaces pour illustrer le travail accompli au cours de notre période de stage. Ces images permettront de visualiser le résultat concret de nos efforts et de mettre en évidence les fonctionnalités développées.

## 2 Nettoyage et collecte des données

### 2.1 La source des données

Nous avons été confrontés à un défi majeur concernant les données fournies par KSF Space initialement sous format Excel. La base de données fournie était de petite taille, ce qui ne suffisait pas à réaliser notre analyse statistique. Afin de résoudre ce problème, nous avons entrepris de trouver des données réelles sur les nanosatellites sur le site web : <https://www.nanosats.eu/database>. Nous avons collecté 72 lignes de données pour chaque année à partir de 2015, ce qui nous a permis d'obtenir un échantillon plus représentatif pour nos analyses et d'obtenir des résultats plus fiables.

### 2.2 Nettoyage des données (cleaning)

Le processus de nettoyage des données était nécessaire pour éliminer les erreurs, les valeurs manquantes et les incohérences présentes dans les données brutes. Pour réaliser ce nettoyage, j'ai utilisé Python et ses bibliothèques dédiées au traitement des données telles que Pandas. J'ai appliqué une série de techniques de nettoyage, notamment la détection et la suppression des doublons, la gestion des valeurs manquantes en les imputant ou en les supprimant selon leur impact sur l'analyse, ainsi que la correction des incohérences ou des valeurs aberrantes.

En effectuant cette étape de nettoyage des données, j'ai pu garantir la qualité et la validité des informations utilisées dans notre projet. Cela a contribué à améliorer la précision et la fiabilité de nos analyses statistiques et de nos indicateurs de performance, renforçant ainsi la pertinence de notre tableau de bord d'aide à la décision.

```

import pandas as pd
df=pd.read_excel('C:\\Users\\king\\Desktop\\INTERNSHIP\\data of nanosatellites.xlsx')
print(df)
print(df.shape)
print(df.columns)
df['lauch_date']=pd.to_datetime(df['lauch_date'],format='mixed')
df['created_at']=pd.to_datetime(df['created_at'],format='mixed')
print(df.info())
df['updated_at']=pd.to_datetime(df['updated_at'],format='mixed')
df['crawl_time'] = pd.to_datetime(df['crawl_time'], format='mixed')
df['last_update_date']=pd.to_datetime(df['last_update_date'],format='mixed')
print(df.info())
print(df.describe(include='object'))
print(df['country'])
print(df.describe())
print(df.isnull().sum())
#df.to_csv('df.csv')

```

Figure 14 : Code de cleaning

### 3.1 Web scraping

Web scraping est une technique utilisée pour extraire automatiquement des données à partir de sites web. Il permet d'extraire des informations structurées à partir de pages web non structurées. Le processus de web scraping implique l'envoi d'une requête HTTP à une page web, l'analyse du contenu HTML de la page et l'extraction des données souhaitées à l'aide de techniques d'analyse et de traitement des données.

Le web scraping est souvent utilisé pour collecter des données à grande échelle, automatiser des tâches de collecte d'informations et alimenter des applications ou des bases de données avec des données en temps réel.

Dans notre situation, nous avons employé la technique de web scraping pour extraire des données relatives aux nanosatellites à partir d'un site web contenant une base de données mondiale dédiée à ces nanosatellites.

```

web.py > ...
1  import json
2  import pandas as pd
3  import requests
4  from bs4 import BeautifulSoup
5  url = "https://www.nanosats.eu/database"
6
7  # Send a GET request to the URL
8  response = requests.get(url)
9  # Check if the request was successful
10 if response.status_code == 200:
11     # Create BeautifulSoup object and specify the parser
12     soup = BeautifulSoup(response.content, 'html.parser')
13     # Find the relevant elements on the page and extract the data
14     # Modify this part according to the structure of the website's HTML
15     column_names = [th.get_text(strip=True) for th in soup.find('tr').find_all('th')]
16     td_elements = soup.find_all('tr')
17     data = []
18     for tr in td_elements:
19         row_data = {}
20         # print(tr.find_all('td'))
21         for td in tr.find_all('td'):
22             print(td.get_text())
23             print('----- \n\n')
24         for i, td in enumerate(tr.find_all('td')):
25             if i == len(column_names) - 1:
26                 photo_td = td.find('a', class_='fancybox-2')
27                 if photo_td:
28                     photo_url = photo_td['href']
29                     row_data[column_names[i]] = "https://www.nanosats.eu/" + photo_url
30             else:
31                 row_data[column_names[i]] = td.get_text(strip=True)
32         data.append(row_data)
33
34     json_data = json.dumps(data, indent=4)
35     print(json_data) # hadi hia data json 7
36     df = pd.DataFrame(data)
37     # Specify the output Excel file path
38     output_file = 'data.xlsx'
39     # Save the DataFrame to an Excel file
40     df.to_excel(output_file, index=False)
41 else:
42     print("Failed to retrieve data from the website.")

```

Figure 15 : Code de scraping



## 3 Les outils de développement utilisés

### 3.1 Visual studio



Figure 16 : Logo de visual studio

Visual Studio Code est un éditeur de code open-source développé par Microsoft qui offre une large gamme de fonctionnalités pour le développement de logiciels. Grâce à son architecture modulaire et à sa prise en charge des extensions, il est capable de prendre en charge une vaste gamme de langages de programmation. Les principales caractéristiques de Visual Studio Code incluent l'auto-complétions, la coloration syntaxique, le débogage et la gestion des commandes git.

### 3.2 My SQL



Figure 17 : Logo de MySQL Workbench

MySQL Workbench est un outil complet et visuel conçu pour faciliter le travail des architectes de bases de données, des développeurs et des administrateurs de bases de données (DBA). Il regroupe plusieurs fonctionnalités essentielles telles que la modélisation des données, le développement SQL et des outils d'administration avancés. Avec MySQL Workbench, il est possible de configurer le serveur, de gérer les utilisateurs, d'effectuer des sauvegardes et bien d'autres tâches liées à la gestion des bases de données. De plus, MySQL Workbench est compatible avec les systèmes d'exploitation Windows, Linux et Mac OS X, offrant ainsi une flexibilité d'utilisation.

### 3.3 HTML



Figure 18 : Logo de HTML

HTML, ou HyperText Markup Language, est un langage de balisage utilisé pour créer et structurer le contenu des pages web. HTML utilise des balises pour marquer différents éléments tels que les titres, les paragraphes, les images, etc.

Ces balises facultatives la façon dont le contenu doit être affiché et permettent aux navigateurs web de comprendre la structure du document et de l'interpréter correctement.

### 3.4 CSS



Figure 19 : Logo de CSS

CSS, ou Cascading Style Sheets, est un langage de feuilles de style utilisé pour décrire la présentation et l'apparence visuelle des documents HTML ou XML. Il permet de contrôler l'aspect des éléments d'une page web, tels que la couleur, la police, la taille, la mise en page, les bordures et les effets visuels.

### 3.5 Java script



Figure 20 : Logo de Javascript

Javascript est un langage de programmation de scripts principalement employé dans les pages web interactives mais aussi pour les serveurs avec l'utilisation (par exemple) de Node.js. C'est un langage orienté objet à prototype, c'est-à-dire que les bases du langage et ses principales interfaces sont fournies par des objets qui ne sont pas des instances de classes, mais qui sont chacun équipés de constructeurs permettant de créer leurs propriétés, et notamment une propriété de prototypage qui permet d'en créer des objets héritiers personnalisés. En outre, les fonctions sont des objets de première classe. Le langage supporte le paradigme objet, impératif et fonctionnel.

### 3.6 Chart.js



Figure 21 : Logo de Chart.js

Chart.js est un logiciel gratuit open source bibliothèque JavaScript pour la visualisation de données, qui prend en charge 8 graphiques types : bar, ligne, zone, tarte (beignet), bulle, radar, polaire, et la dispersion. Créé par Londres-Basé développeur web Nick Downie en 2013.

### 3.7 Bootstrap



Figure 22 : Logo de bootstrap

Bootstrap est une collection d'outils utiles à la création du design (graphisme, animation et interactions avec la page dans le navigateur, etc.) de sites et d'applications web. C'est un ensemble qui contient des codes HTML et CSS, des formulaires, boutons, outils de navigation et autres éléments interactifs, ainsi que des extensions JavaScript en option. C'est l'un des projets les plus populaires sur la plate-forme de gestion de développement GitHub.

### 3.8 Python



Figure 23 : Logo de python

Python est un langage de programmation interprété, multiparadigme et multiplateformes. Il favorise la programmation impérative structurée, fonctionnelle et orientée objet. Flask est un micro framework open-source de développement web en Python. Il est classé comme microframework car il est très léger. Flask a pour objectif de garder un noyau simple mais extensible.

### 3.9 Flask



Figure 24 : Logo de flask

Flask est un micro framework open-source de développement web en Python. Il est classé comme microframework car il est très léger. Flask a pour objectif de garder un noyau simple mais extensible.

### 3.10 SQL



SQL est un langage informatique normalisé servant à exploiter des bases de données relationnelles. La partie langage de manipulation des données de SQL permet de rechercher, d'ajouter, de modifier ou de supprimer des données dans les bases de données relationnelles.

Figure 25 : Logo de SQL

## 4 Les interfaces graphiques de l'application

Dans cette partie, nous allons présenter en détail les interfaces graphiques de notre application, en mettant l'accent sur leur fonctionnement et leur rôle dans l'expérience utilisateur.

Nous décrirons aussi comment ces interfaces permettent aux utilisateurs d'interagir avec l'application de manière efficace.

### 4.1 Authentification

Cette interface offre un aperçu de la procédure d'authentification, où l'utilisateur est invité à fournir leurs informations d'identification. Une fois que l'authentification est réussie, l'utilisateur est dirigé vers l'interface d'accueil, où ils ont accès à toutes les fonctionnalités de l'application.

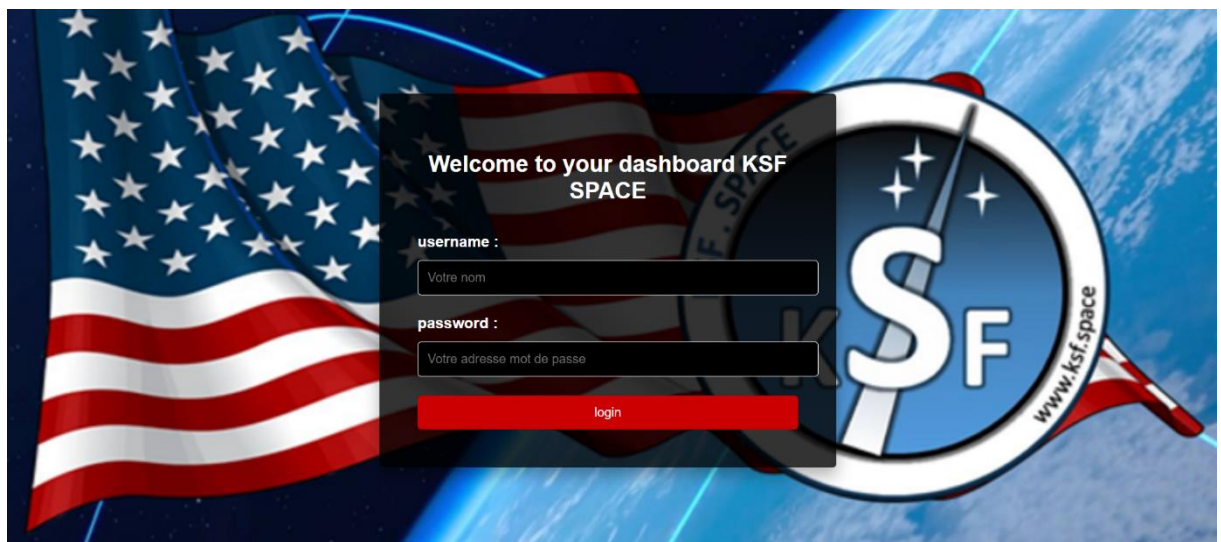


Figure 26 : Authentification

## 4.2 Page d'accueil

La page d'accueil représente un menu regroupant tous les indicateurs pertinents pour les nanosatellites, ce qui offre à l'utilisateur la possibilité de visualiser et de filtrer les données en fonction de ses besoins spécifiques. Cela permet une expérience personnalisée, offrant une flexibilité dans l'exploration et l'analyse des données liées aux nanosatellites.

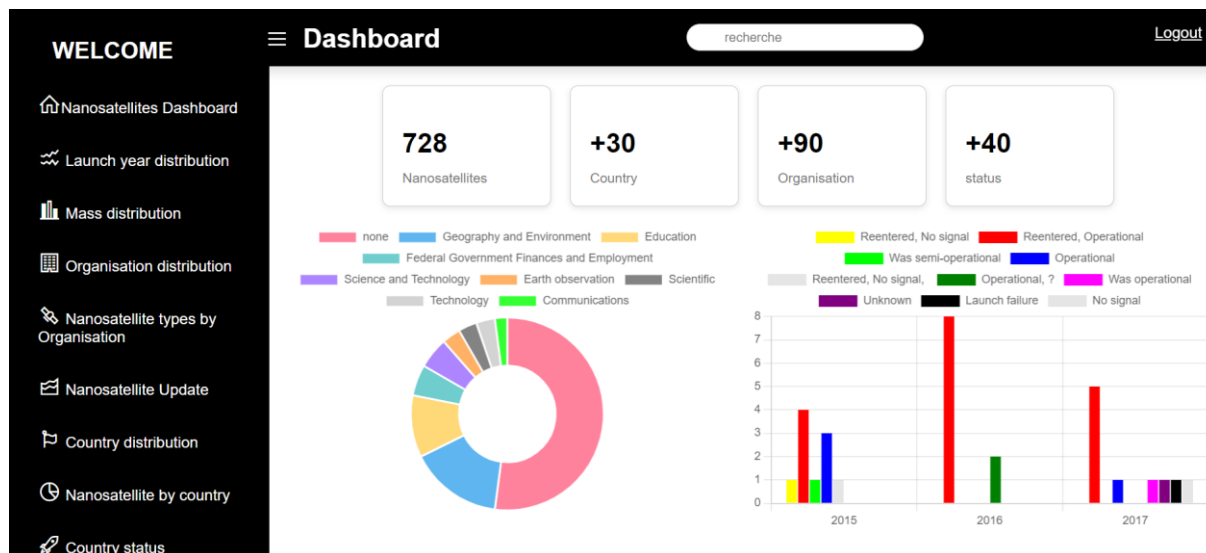


Figure 27 : Page d'accueil

## 4.3 Le premier Indicateur

Cette interface permet de visualiser la distribution des nanosatellites lancés chaque année. Il offre la possibilité de filtrer les années selon les besoins afin d'obtenir une vue précise des nanosatellites lancés chaque année. Cela permet de prendre des décisions éclairées concernant l'état de ces nanosatellites, en utilisant cette fonctionnalité, il est possible d'analyser les tendances et les évolutions dans le domaine des nanosatellites au fil du temps.

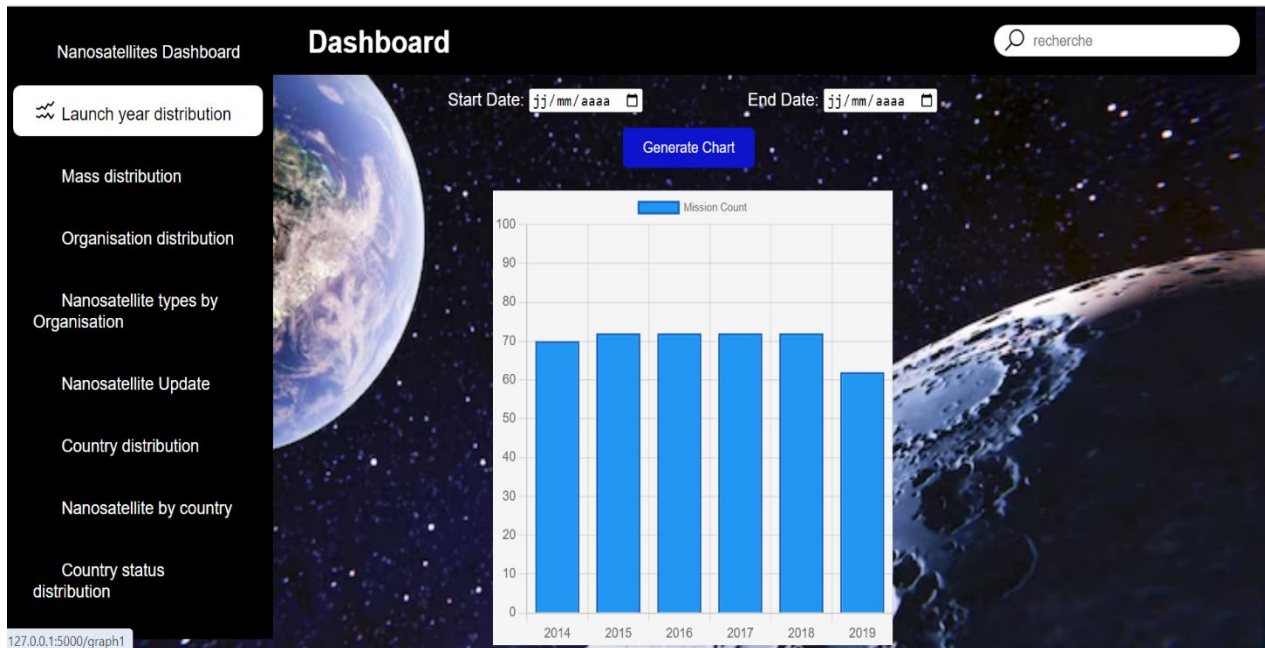


Figure 28 : Indicateur1 (DNL)

#### 4.4 Le deuxième indicateur

Ce graphique représente la somme totale des masses des nanosatellites pour une période d'années indiquée par l'utilisateur. Il permet d'obtenir une vision claire de la masse dominante, dans ce cas est de 3U. Cette visualisation permet de mieux comprendre la répartition des masses des nanosatellites sur la période choisie, en mettant en évidence la masse la plus fréquente ou dominante. Cela peut être utile pour analyser les caractéristiques générales des nanosatellites en fonction de leur masse.

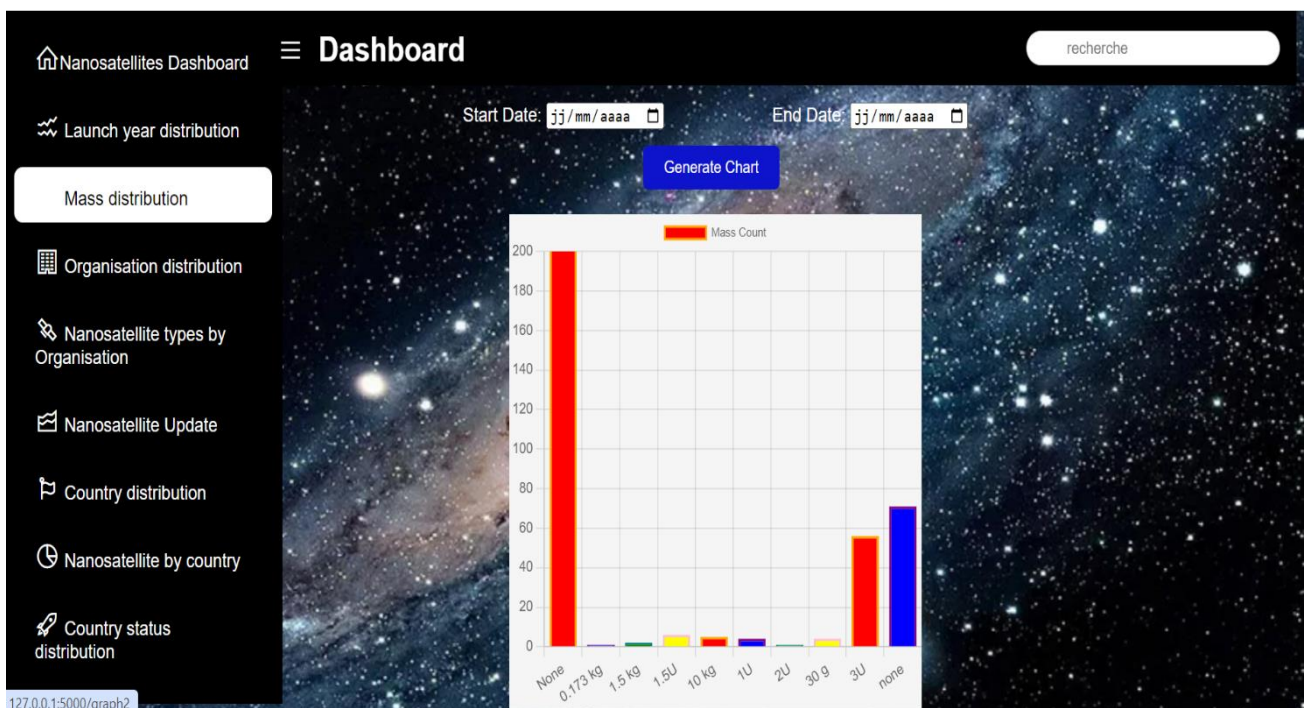


Figure 29 : Indicateur2 (DM)



## 4.5 Le troisième indicateur

Ce graphique présente la répartition des organisations par une période d'année et met en évidence les pays qui comptent le plus grand nombre d'organisations. Il permet d'obtenir une vue claire des organisations présente chaque année, en mettant l'accent sur les pays qui abritent le plus grand nombre d'organisations.

Cette visualisation facilite l'identification des pays dominants en termes d'organisations et permet d'analyser les tendances et les variations dans la répartition géographique des organisations au fil du temps. Cela peut être utile pour comprendre la dynamique des organisations à l'échelle mondiale.

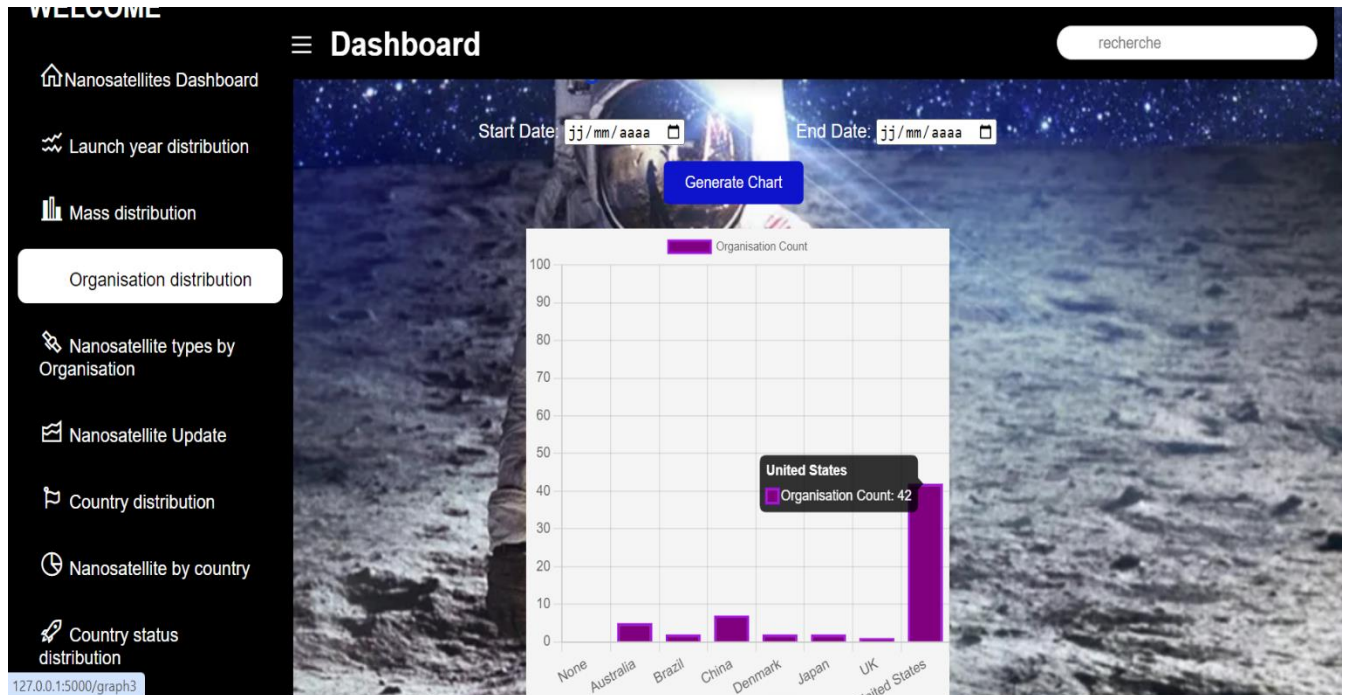


Figure 30 : Indicateur3 (DO)

## 4.6 Le quatrième indicateur

Cette interface permet à l'utilisateur d'obtenir une vision claire du nombre de catégories présentes dans l'organisation saisie et les types aussi. De plus, elle affiche la catégorie de nanosatellite la plus utilisée par cette organisation.

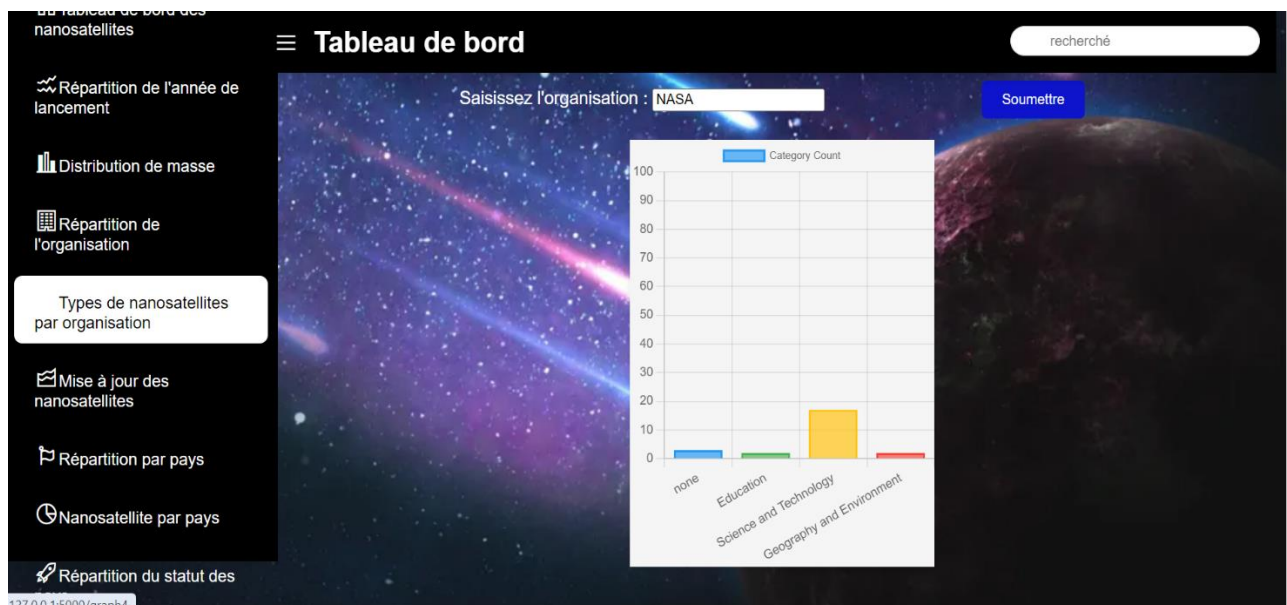


Figure 31 : Indicateur4 (TNO)

## 4.7 Le cinquième indicateur

Ce graphique présente l'évolution du nombre de nanosatellites mises à jour au fil des années ce qui nous permet d'identifier l'année qui a le plus grand nombre de nanosatellites mises à jour, donnant ainsi une idée générale de cette tendance.

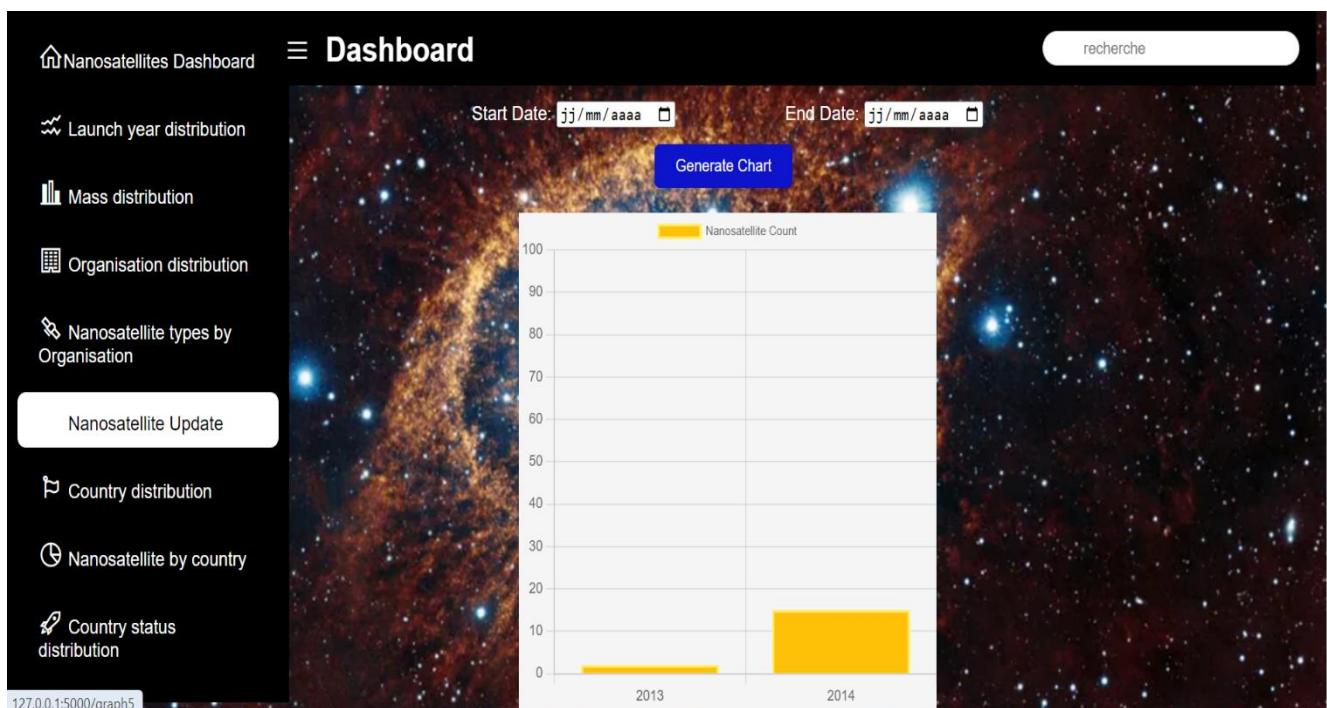


Figure 32 : Indicateur5 (MJN)



## 4.8 Le seizième indicateur

Ce graphique présente le nombre de nanosatellites dans chaque pays, ce qui nous donne une répartition des différentes nations présentées dans nos données.

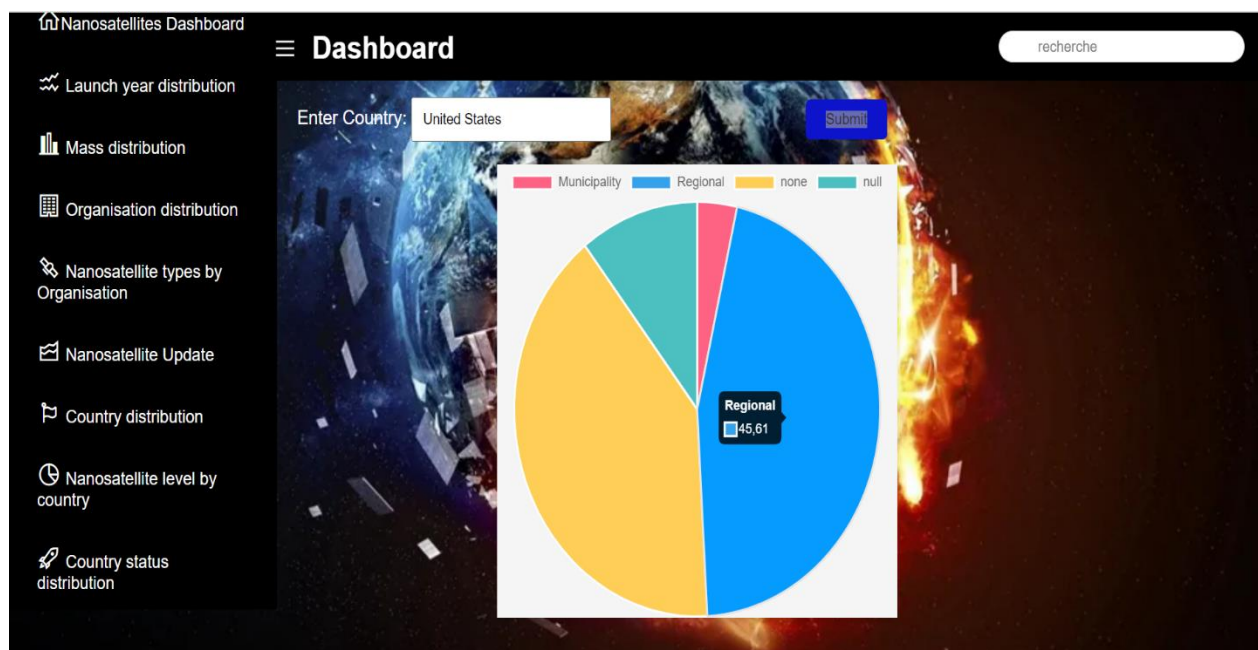


Figure 33 : Indicateur6 (DN)

## 4.9 Le septième indicateur

Ce graphique illustre le pourcentage des nanosatellites classés selon leur niveau de mission, que ce soit au niveau national, régional, ou autre, en fonction du pays saisi, dans cet exemple nous avons choisi le pays United states pour savoir sa répartition de niveau.

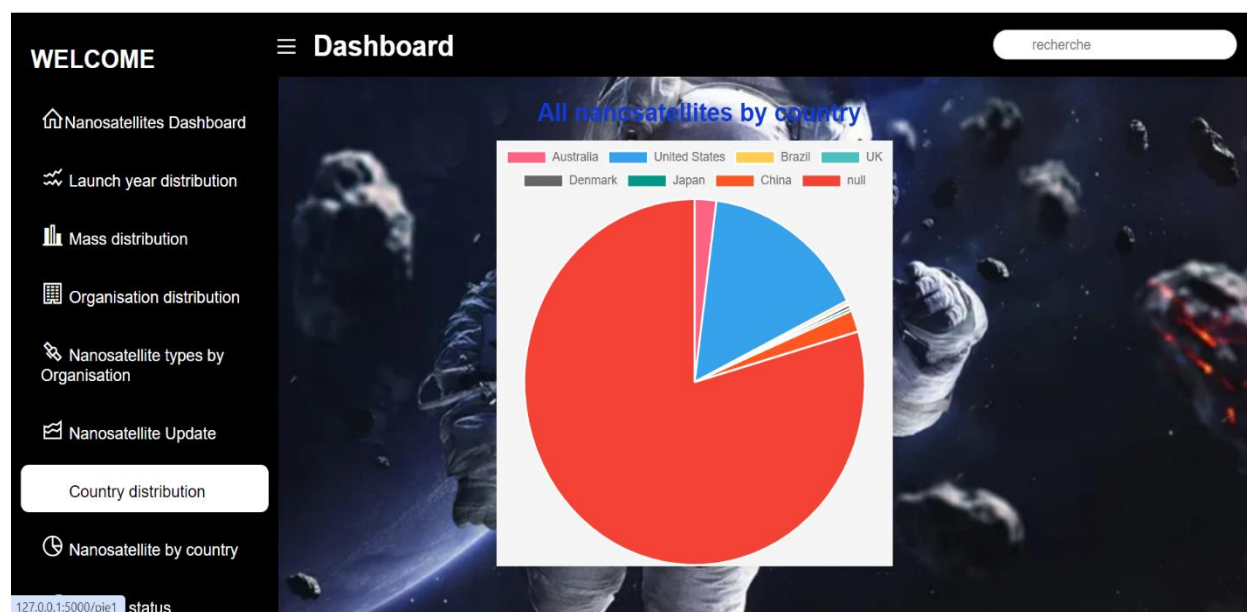


Figure 34 : Indicateur7 (PNN)

#### 4.10 Le huitième indicateur

Cette interface affiche la répartition des statuts des nanosatellites (opérationnel, semi-opérationnel...) en fonction du pays saisi par l'utilisateur.

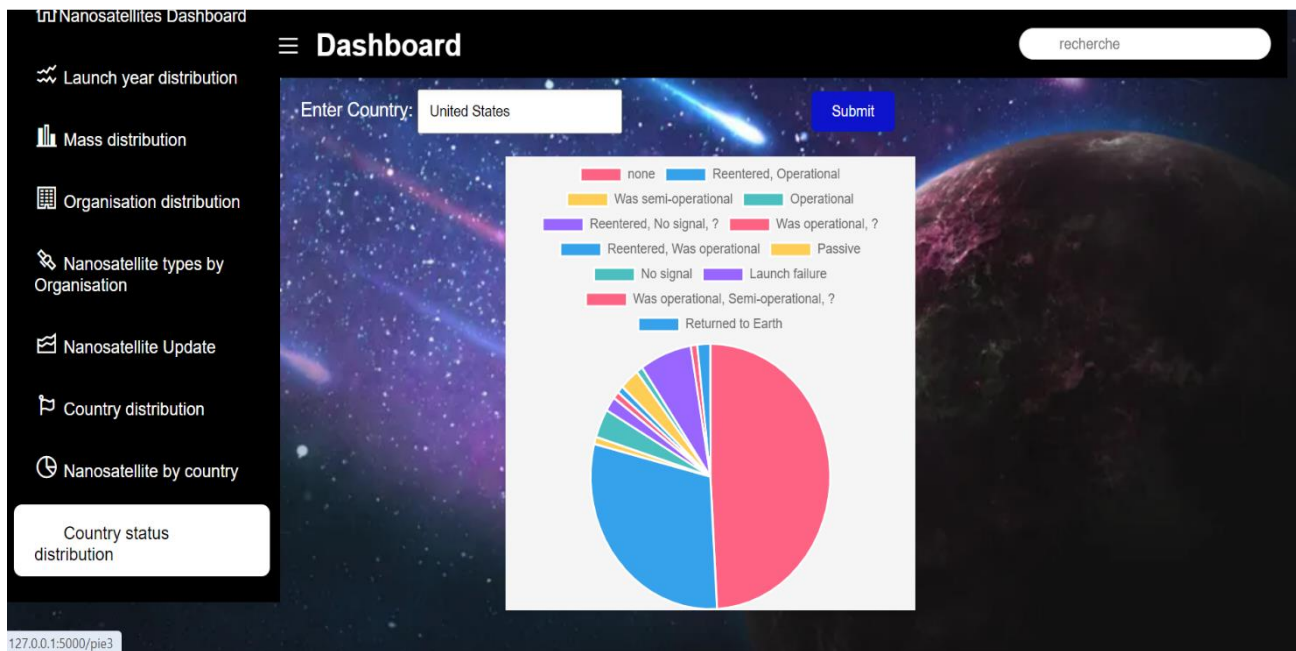


Figure 35 : Indicateur8 (RSN)

## Chapitre 4 : Analyse statistique des Nanosatellites

# 1 Introduction

Dans ce chapitre, on va faire une analyse de données en commençant par une analyse préliminaire de données à l'aide de statistiques descriptives. Cette analyse nous permettra de comprendre les caractéristiques principales des données et de les visualiser à l'aide de graphiques et de tableaux. Nous examinons les pourcentages et les fréquences pour chaque variable afin d'obtenir une vue d'ensemble des différentes catégories présentes dans les données. Cela nous permet d'identifier les relations potentielles entre les variables, ce qui constitue une base solide pour l'analyse plus approfondie des données.

## 2 Etude préliminaire de données

La statistique descriptive décrit et résume les caractéristiques essentielles d'un ensemble de données. Elle fournit des outils et des techniques permettant de présenter et d'analyser les données de manière informative et concise.

L'objectif de la statistique descriptive est de synthétiser les informations contenues dans les données, de les organiser de manière significative et de les rendre plus compréhensibles.

### 2.1 Tableau des variables quantitatives

Ce tableau calcule diverses mesures statistiques spécifiquement pour les variables catégorielles.

- **Count** : Cela représente le nombre de valeurs non nulles dans chaque colonne. Il indique le nombre d'observations disponibles pour l'analyse.
- **Unique** : indique le nombre de catégories uniques ou de valeurs distinctes présentes dans chaque colonne.
- **Top** : Cela affiche la catégorie ou la valeur la plus fréquente dans chaque colonne.
- **Freq** : Cela indique la fréquence ou le nombre de la catégorie ou de la valeur supérieure dans chaque colonne.

	country	level	language	catalog_type	nanosatellite_name	mission_name	english_keywords	english_category	english_description	laucnh_date	...	created_at
count	728	728	728	728	728	728	728	728	728	728	...	71
unique	38	4	1	3	658	40	76	12	240	129	...	71
top	United States	none	English	none	none	none	Earth observation	Earth observation	Constellation for optical Earth observation wi...	2023-01-03 00:00:00	...	2014-12-04 00:30:13.216967
freq	559	657	728	657	71	657	495	393	321	73	...	1

4 rows x 22 columns

Figure 36 : Tableau des variables quantitatives

Par exemple le pays le plus dominant de création et de lancement des nanosatellites est "United States", et la catégorie des nanosatellites la plus fréquente est "Earth observation".

### 2.1.1 Variable quantitative : country

Dans cette variable, plusieurs pays sont impliqués dans le lancement de nanosatellites, comme le montre l'image et le graphique. Certains pays sont répertoriés plusieurs fois, ce qui indique qu'ils ont effectué plusieurs lancements. Cependant, on observe que les États-Unis est le pays le plus dominant en termes de nombre de lancements de nanosatellites, suivi de près par la Chine et l'Australie ainsi de suite. Cette distribution des pays induit une concentration de l'activité de lancement de nanosatellites dans ces régions spécifiques.

United States	559
China	32
Australia	16
UK	13
Japan	12
Switzerland	12
Spain	10
Canada	7
South Korea	6
France	6
Germany	6
Denmark	5
Israel	4
Turkey	3
Poland	3
Lithuania	2
Taiwan	2
United Arab Emirates	2
Norway	2
Italy	2
Bulgaria	2
Brazil	2
Belgium	2
Argentina	2
...	
Kazakhstan	1
Costa Rica	1

Figure 37 : la frequence de la variable country

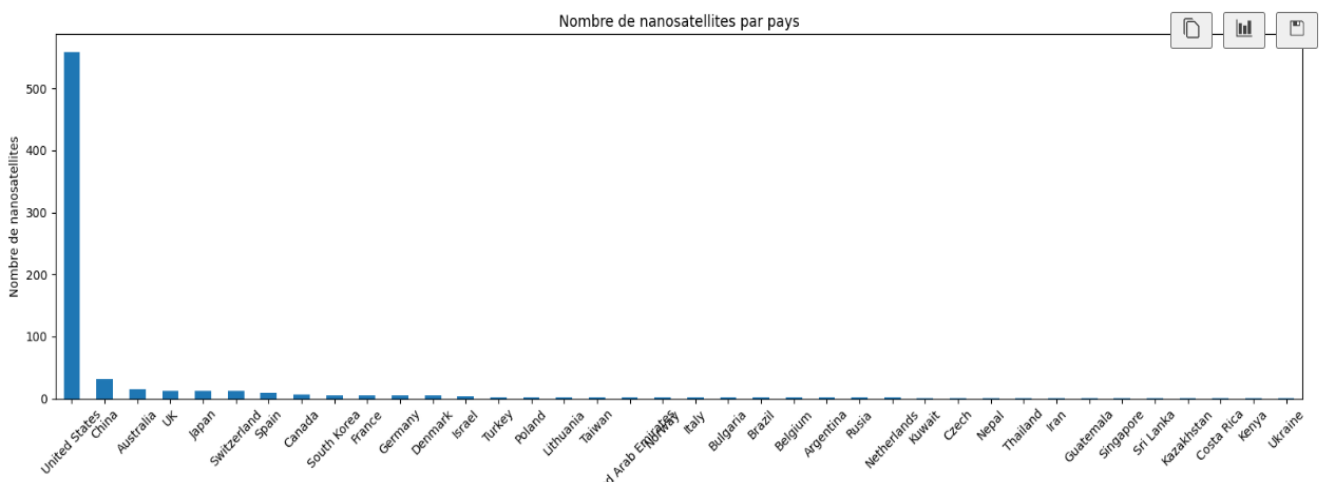


Figure 38 : Distribution de nanosatellites par pays

### 2.1.2 Variable quantitative : level

Ce graphique nous permet de visualiser le nombre de nanosatellites dans chaque niveau de classification. Cela fournit une représentation visuelle claire de la répartition des nanosatellites par niveau de classification, et on voit aussi que la plupart des nanosatellites n'ont pas un niveau, ensuite on a le niveau régional qui a un nombre de Nanos inférieur à cent, puis le niveau national et le niveau municipalité est presque nul.

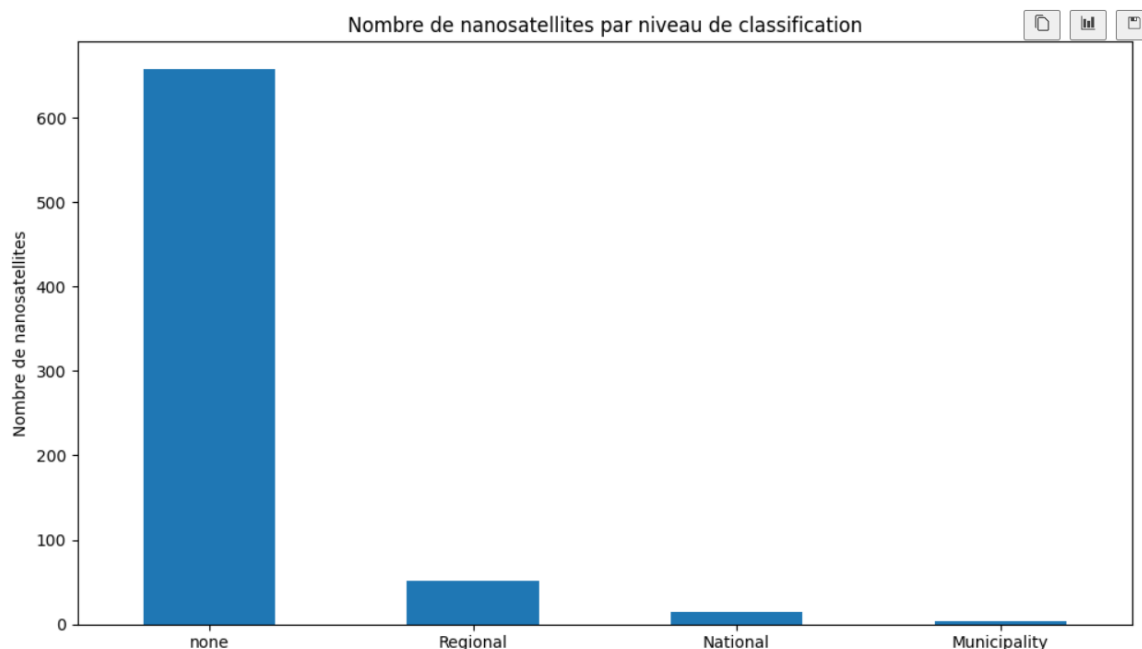
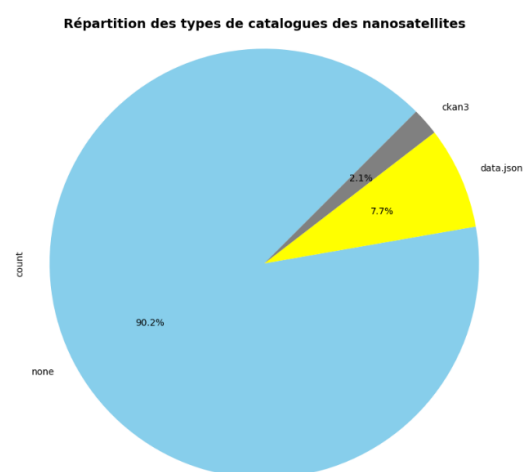


Figure 39 : Répartition de nanosatellites par niveau

### 2.1.3 Variable quantitative : catalogue

Ce diagramme circulaire permet de visualiser rapidement la répartition des types de catalogue des nanosatellites comme suit :



- 90% des nanosatellites n'ont pas un catalogue spécifique, cela signifie qu'ils ne sont pas associés à un type de catalogue particulier.

- Environ 7% des nanosatellites sont associés au type de catalogue data.json

- 2% des nanosatellites sont associés au type ckan

Ce graphique donne une idée de la proportion de chaque type de catalogue dans l'ensemble des nanosatellites étudiés.

Figure 40 : Répartition des types de catalogue des nanosatellites

### 2.1.4 Variable qualitative : catégorie

Ce graphique représente la répartition des catégories de nanosatellites par pays .il montre les différentes catégories dans chaque pays, par exemple Australie n'a pas aucune catégorie par contre aux États-Unis on observe différentes catégories de nanosatellites telles que la technologie, la communication, l'éducation, la science, etc.

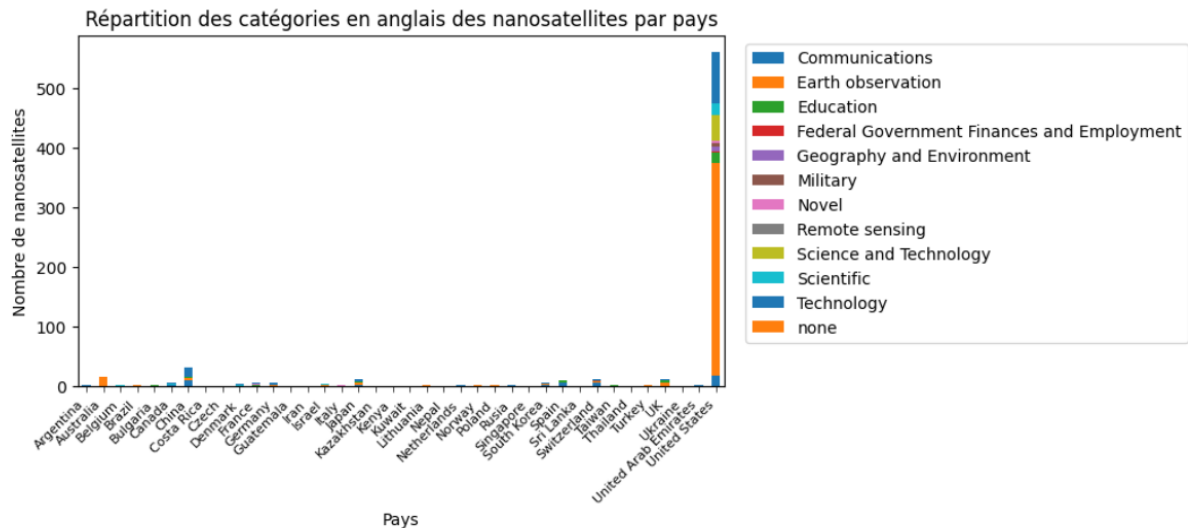


Figure 41 : Répartition des catégories des nanosatellites par pays

### 2.1.5 Variable quantitative : licence

Ce graphique illustre la répartition des licences des nanosatellites et il montre que la plupart des nanosatellites ne possèdent pas une licence tandis que de 2% qui possèdent une licence.

Cette visualisation permet de constater clairement que la plupart des nanosatellites fonctionnent sans une licence spécifique, ce qui peut être dû à diverses raisons, telles que des réglementations moins strictes pour les nanosatellites ou des situations particulières dans lesquelles les licences ne sont pas nécessaires.

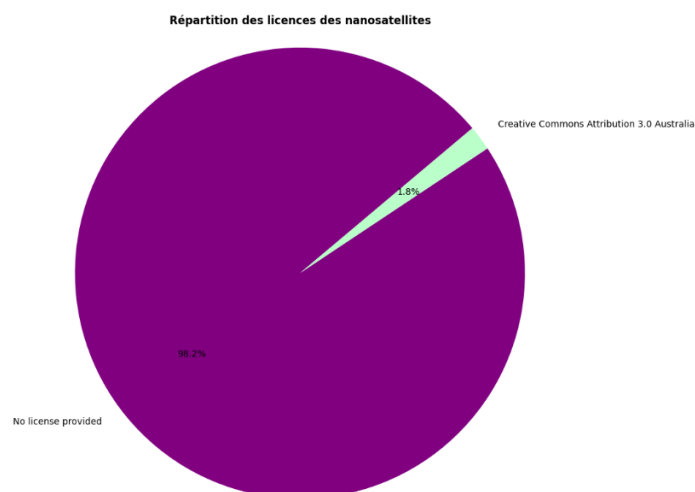


Figure 42 : Répartition des licences des nanosatellites

### 3 Analyse factorielle des correspondances

L'analyse factorielle des correspondances est une méthode statistique utilisée pour analyser des données qualitatives. Elle vise à identifier les associations et les relations entre les différentes catégories de variables qualitatives dans un tableau croisé. L'AFC permet de réduire la dimensionnalité des données et de les représenter graphiquement dans un espace de dimensions réduites.

Cela facilite l'interprétation visuelle des relations et des structures sous-jacentes dans les données catégorielles. L'AFC est couramment utilisé dans divers domaines tels que la sociologie, le marketing, la linguistique, etc., pour explorer et analyser les données qualitatives de manière exploratoire.

#### 3.1 Table de contingence

La table de contingence montre le nombre de satellites nanosatellites dans différentes catégories anglaises pour chaque pays.

- Pour chaque pays, on a plusieurs catégories (Communications, Earth observation, Education, Science, etc.).
- Dans chaque cellule de la table, on a le nombre des nanosatellites correspondant à une catégorie anglaise spécifique et à un pays spécifique.



...	english_category	Communications	Earth observation	Education	\
	country				
	Argentina	0	0	0	
	Australia	0	0	0	
	Belgium	1	0	0	
	Brazil	1	1	0	
	Bulgaria	0	0	2	
	Canada	2	1	0	
	China	11	4	2	
	Costa Rica	0	0	1	
	Czech	0	1	0	
	Denmark	0	1	0	
	France	0	1	1	
	Germany	0	2	0	
	Guatemala	0	0	0	
	Iran	0	0	0	
	Israel	0	2	0	
	Italy	0	0	0	
	Japan	2	4	3	
	Kazakhstan	0	0	1	
	Kenya	0	0	0	
	Kuwait	0	1	0	
	Lithuania	1	1	0	
	Nepal	0	0	1	
	Netherlands	0	0	0	
...					
	UK	2	0		
	Ukraine	0	0		
	United Arab Emirates	1	0		
	United States	86	0		

Figure 43 : Table de contingence

La table de contingence présente les fréquences des différentes catégories pour chaque pays, comme suit :

- Communications : On peut observer que les pays suivants ont des nanosatellites liés à la communication : Belgique (1), Brésil (1), Canada (2), Chine (11), Danemark (0), France (0), Allemagne (0), Israël (0), Japon (2), Lituanie (1), Pays-Bas (0), Royaume-Uni (2), Émirats arabes unis (1) et États-Unis (86).
- Earth observation : Certains pays ont des nanosatellites dédiés à l'observation de la Terre. Les pays concernés sont le Brésil (1), le Canada (1), la Chine (4), la République tchèque (1), le Danemark (1), la France (1), l'Allemagne (2), Israël (2), le Japon (4), le Koweït (1) et la Lituanie (1).
- Education : Certains pays ont des nanosatellites liés à l'éducation. Les pays concernés sont la Bulgarie (2), le Costa Rica (1), la République tchèque (0), le Danemark (0), la France (1), l'Allemagne (0), Israël (0), le Japon (3), le Kazakhstan (1), le Kenya (0), le Koweït (0), la Lituanie (0) et le Népal (1).
- Science : Certains pays ont des nanosatellites liés à la science. Les pays concernés sont la Bulgarie (0), le Costa Rica (0), la République tchèque (1), le Danemark (0), la France (0), l'Allemagne (0), Israël (0), le Japon (3), le Kazakhstan (1), le Kenya (0), le Koweït (0), la Lituanie (1) et le Népal (0).
- Autres catégories : Il existe également d'autres catégories de nanosatellites.

Ces résultats nous permettent d'identifier les pays qui sont actifs dans différentes catégories de nanosatellites.

## 3.2 Table profil-ligne

La table de profil ligne représente la répartition des différentes catégories (Communications, Earth observation, Education, etc.) pour chaque pays. Chaque valeur dans la table indique la proportion de la catégorie respective par rapport à la somme de toutes les catégories pour le pays donné.

Par exemple, prenons la ligne pour le pays "Argentine". Toutes les valeurs de catégorie pour ce pays sont nulles (0,0), ce qui signifie qu'il n'y a aucune observation dans ces catégories pour l'Argentine.

En revanche, prenons la ligne pour le pays "Belgique". On peut observer que la catégorie "Communications" représente 50% des observations pour la Belgique, tandis que les autres catégories ont une proportion nulle (0,0).

De même, pour le pays "Chine", la catégorie "Communications" représente environ 34,4% des observations, la catégorie "Earth observation" représente environ 12,5%, la catégorie "Education" représente environ 6,2%, et ainsi de suite pour les autres catégories.

La table de profil ligne nous permet de voir comment les différentes catégories sont réparties pour chaque pays, en termes de proportion par rapport à l'ensemble des catégories pour ce pays.

english_category	Communications	Earth observation	Education
country			
Argentina	0.000000	0.000000	0.000000
Australia	0.000000	0.000000	0.000000
Belgium	0.500000	0.000000	0.000000
Brazil	0.500000	0.500000	0.000000
Bulgaria	0.000000	0.000000	1.000000
Canada	0.285714	0.142857	0.000000
China	0.343750	0.125000	0.062500
Costa Rica	0.000000	0.000000	1.000000
Czech	0.000000	1.000000	0.000000
Denmark	0.000000	0.200000	0.000000
France	0.000000	0.166667	0.166667
Germany	0.000000	0.333333	0.000000
Guatemala	0.000000	0.000000	0.000000
Iran	0.000000	0.000000	0.000000
Israel	0.000000	0.500000	0.000000
Italy	0.000000	0.000000	0.000000
Japan	0.166667	0.333333	0.250000
Kazakhstan	0.000000	0.000000	1.000000
Kenya	0.000000	0.000000	0.000000
Kuwait	0.000000	1.000000	0.000000
Lithuania	0.500000	0.500000	0.000000
Nepal	0.000000	0.000000	1.000000
Netherlands	0.000000	0.000000	0.000000
...			
UK	0.153846	0.0000	
Ukraine	0.000000	0.0000	
United Arab Emirates	0.500000	0.0000	
United States	0.153846	0.0000	

Figure 44 : Table profil-ligne

### 3.3 Table profil-colonne

D'après la table de profil des colonnes ci-dessous, voici les observations pour chaque catégorie par rapport aux différents pays :

- **Communications** : La Chine présente la plus grande proportion d'observations dans cette catégorie (environ 20,4 %), suivie des États-Unis (environ 66,2 %). La plupart des autres pays ont une proportion très faible ou aucune observation dans cette catégorie.
- **Observation de la Terre** : La Chine présente une proportion relativement élevée d'observations dans cette catégorie (environ 1,0 %), suivie du Japon (environ 1,0 %) et des États-Unis (environ 0,6 %).  
Les autres pays ont généralement une proportion faible ou aucune observation dans cette catégorie.
- **Éducation** : La Chine présente la plus grande proportion d'observations dans cette catégorie (environ 5,4 %), suivie du Japon (environ 8,1 %). Plusieurs autres pays, tels que la Bulgarie, le Canada et la France, ont également une proportion non nulle d'observations dans cette catégorie, bien que relativement faible par rapport à la Chine et au Japon.
- **Énergie** : Les États-Unis présentent la plus grande proportion d'observations dans cette catégorie (environ 10,8 %), suivis du Japon (environ 4,2 %) et de la Chine (environ 3,8 %).  
Les autres pays ont généralement une proportion inférieure ou aucune observation dans cette catégorie.
- **Finance** : Les États-Unis présentent de loin la plus grande proportion d'observations dans cette catégorie (environ 98,4 %), indiquant une forte domination dans les observations liées aux finances. Les autres pays ont des proportions négligeables ou aucune observation dans cette catégorie.
- **Santé** : Le Japon présente la plus grande proportion d'observations dans cette catégorie (environ 5,8 %), suivi des États-Unis (environ 2,5 %) et de la Chine (environ 1,6 %). Les autres pays ont généralement une proportion inférieure ou aucune observation dans cette catégorie.
- **Transport** : Les États-Unis présentent la plus grande proportion d'observations dans cette catégorie (environ 8,5 %), suivis du Japon (environ 4,2 %) et de la Chine (environ 1,9 %). Les autres pays ont généralement une proportion inférieure ou aucune observation dans cette catégorie.

...	english_category	Communications	Earth observation	Education	\
	country				
	Argentina	0.000000	0.000000	0.000000	
	Australia	0.000000	0.000000	0.000000	
	Belgium	0.018519	0.000000	0.000000	
	Brazil	0.018519	0.002545	0.000000	
	Bulgaria	0.000000	0.000000	0.054054	
	Canada	0.037037	0.002545	0.000000	
	China	0.203704	0.010178	0.054054	
	Costa Rica	0.000000	0.000000	0.027027	
	Czech	0.000000	0.002545	0.000000	
	Denmark	0.000000	0.002545	0.000000	
	France	0.000000	0.002545	0.027027	
	Germany	0.000000	0.005089	0.000000	
	Guatemala	0.000000	0.000000	0.000000	
	Iran	0.000000	0.000000	0.000000	
	Israel	0.000000	0.005089	0.000000	
	Italy	0.000000	0.000000	0.000000	
	Japan	0.037037	0.010178	0.081081	
	Kazakhstan	0.000000	0.000000	0.027027	
	Kenya	0.000000	0.000000	0.000000	
	Kuwait	0.000000	0.002545	0.000000	
	Lithuania	0.018519	0.002545	0.000000	
	Nepal	0.000000	0.000000	0.027027	
	Netherlands	0.000000	0.000000	0.000000	
...					
	UK	0.015385	0.0		
	Ukraine	0.000000	0.0		
	United Arab Emirates	0.007692	0.0		
	United States	0.661538	0.0		

Figure 45 : Table profil-colonne

### 3.4 Test de KHI-DEUX

Les résultats du test du khi-deux sont les suivants :

Distance du khi-deux : 1616.2415778687737

Valeur p : 1.8521124896102137e-143 (très proche de zéro)

Degrés de liberté : 407

```

chi2_distance, p_value, dof, expected = chi2_contingency(contingency_table)

# Print the chi-square distance
print("Chi-square distance:", chi2_distance)

# Print the p-value
print("p-value:", p_value)

# Print the degrees of freedom
print("Degrees of freedom:", dof)

```

[19] Python

```

... Chi-square distance: 1616.2415778687737
p-value: 1.8521124896102137e-143
Degrees of freedom: 407

```

Figure 46 : Test KHI-DEUX

La distance du khi-deux mesure la différence entre les valeurs observées et les valeurs attendues dans le tableau de contingence. Une distance plus élevée indique une plus grande différence entre les valeurs observées et attendues.

La valeur p est extrêmement faible, ce qui suggère que les différences observées dans le tableau de contingence ne sont pas simplement le résultat du hasard. En d'autres termes, il y a une association significative entre les variables représentées dans le tableau.

Les degrés de liberté représentent le nombre de catégories indépendantes dans le tableau de contingence. Dans ce cas, il y a 407 degrés de liberté, ce qui correspond au nombre total de cellules du tableau moins le nombre de lignes et de colonnes

### 3.5 Matrice d'inertie

La matrice d'inertie semble représenter une mesure de similarité ou de dissimilarité entre les points d'un ensemble de données. Chaque élément de la matrice représente la distance entre deux points.

Les éléments diagonaux de la matrice (lorsque l'indice de ligne est égal à l'indice de colonne) sont tous nuls, ce qui est attendu car ils représentent la distance ou la dissimilarité entre un point et lui-même, qui est toujours zéro. Les autres éléments de la matrice représentent la distance ou la dissimilarité entre les différents points. Les valeurs peuvent varier de zéro pour les points identiques à des valeurs plus élevées pour les points plus éloignés ou plus dissimilaires.

```
.. Inertia Matrix:
[[0.      1.      1.      ... 1.      1.      1.      ]
 [1.      0.      1.      ... 1.      1.      1.      ]
 [1.      1.      0.      ... 0.25    0.44444444 1.      ]
 ...
 [1.      1.      0.25    ... 0.      0.25    1.      ]
 [1.      1.      0.44444444 ... 0.25    0.      1.      ]
 [1.      1.      1.      ... 1.      1.      0.      ]]
```

Figure 47 : Matrice d'inertie

### 3.6 Les valeurs propres

Les valeurs propres représentent les valeurs propres de la matrice d'inertie. Chaque valeur propre est un nombre complexe avec une partie réelle et une partie imaginaire représentée par "j". Voici une traduction des valeurs propres :

```

... Eigenvalues:
[ 3.24086042e+01+0.0000000e+00j -5.53252208e+00+0.0000000e+00j
-4.73490931e+00+0.0000000e+00j -4.56013169e+00+0.0000000e+00j
-3.06167024e+00+0.0000000e+00j -2.71448398e+00+0.0000000e+00j
-2.31345994e+00+0.0000000e+00j -1.96969464e+00+0.0000000e+00j
-1.50137731e+00+0.0000000e+00j -1.24584117e+00+0.0000000e+00j
-1.09564058e+00+0.0000000e+00j -1.02583477e+00+0.0000000e+00j
-6.62084640e-01+0.0000000e+00j -5.20515072e-01+0.0000000e+00j
-3.95022093e-01+0.0000000e+00j 2.93554579e-01+0.0000000e+00j
2.42241508e-01+0.0000000e+00j -3.36479174e-01+0.0000000e+00j
-3.08339809e-01+0.0000000e+00j -2.44949485e-01+0.0000000e+00j
-1.85641816e-01+0.0000000e+00j 1.83185869e-01+0.0000000e+00j
-8.45888377e-02+0.0000000e+00j 1.22167692e-01+0.0000000e+00j
1.05343494e-01+0.0000000e+00j 8.94650248e-02+0.0000000e+00j
5.90518102e-02+0.0000000e+00j -1.04275075e-02+0.0000000e+00j
-1.00000000e+00+0.0000000e+00j 2.88590400e-16+0.0000000e+00j
-2.11048087e-16+0.0000000e+00j 1.59727254e-16+0.0000000e+00j
-5.64127158e-17+6.3109857e-17j -5.64127158e-17-6.3109857e-17j
2.14566604e-17+1.3242903e-17j 2.14566604e-17-1.3242903e-17j
-3.58559615e-17+0.0000000e+00j -1.22945491e-17+0.0000000e+00j]

```

Figure 48 : Valeurs propres

### 3.7 Cordonnée de la ligne AFC

Dans ce graphique on l'AFC en terme ligne on remarque qu'Australie plus éloignée par rapport aux autre pays ce qui signifie que l'Australie est différente de toute les pays des nanosatellites, et on observe que pays telles que UK , canada , South Korea , Japan , United Arab Emirates et Switzer land sont très proches entre eux ce qui montre qu'ils ont presque la même catégorie des nanosatellites, par contre le pays Israël est très loin de ces pays donc il a une catégorie différente de ces pays.

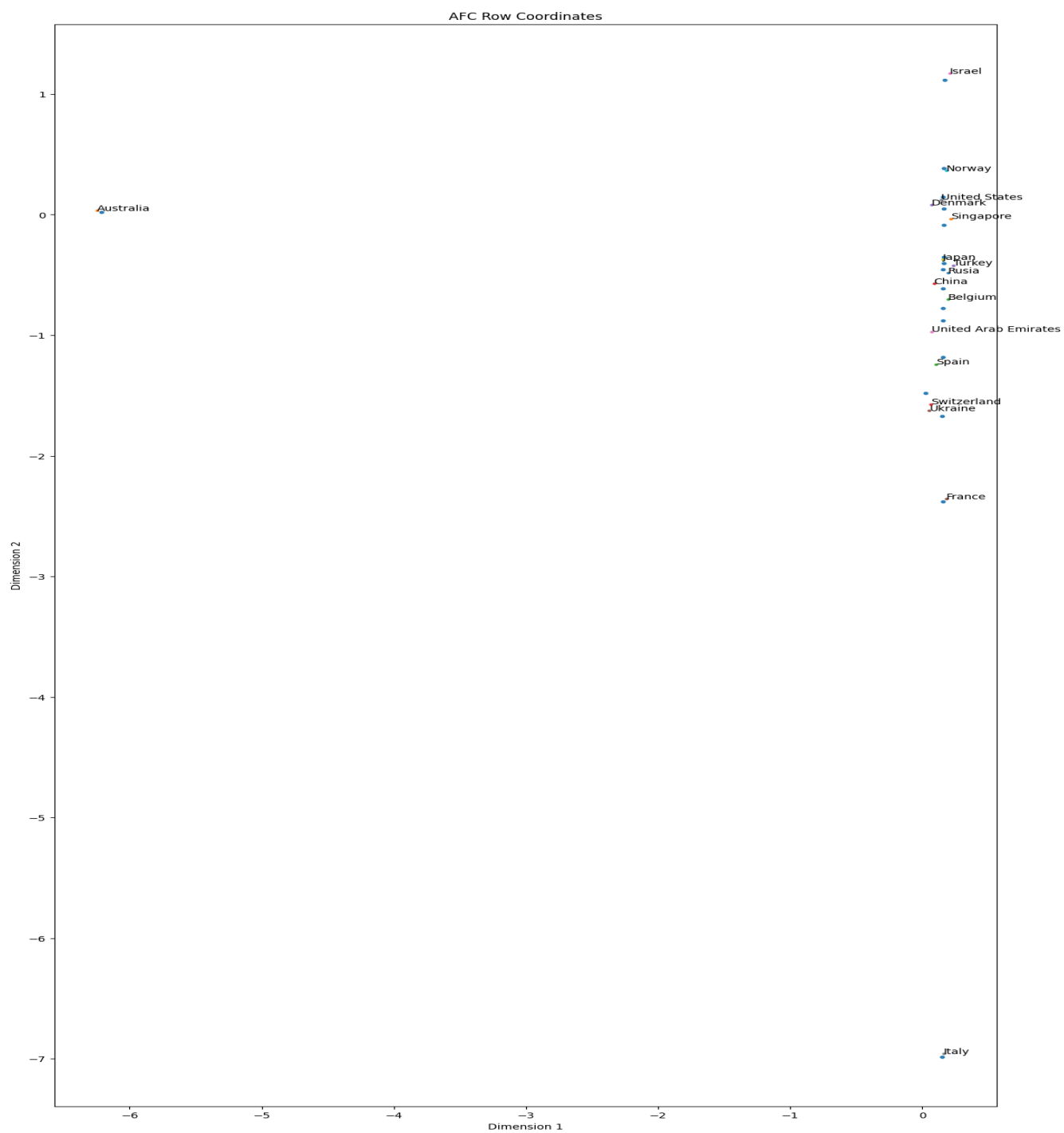


Figure 49 : Coordonnées de la ligne AFC

### 3.8 Cordonnée de la colonne AFC

Dans ce graphique on remarque qu'il existe des nanosatellites qui n'ont pas une catégorie spécifique ce qui correspond au point none ce qui signifie le pays Australie et la catégorie remote sensing correspond à Israël, par contre la majorité des nanosatellites ont les mêmes catégories telles que science, earth observation, etc.

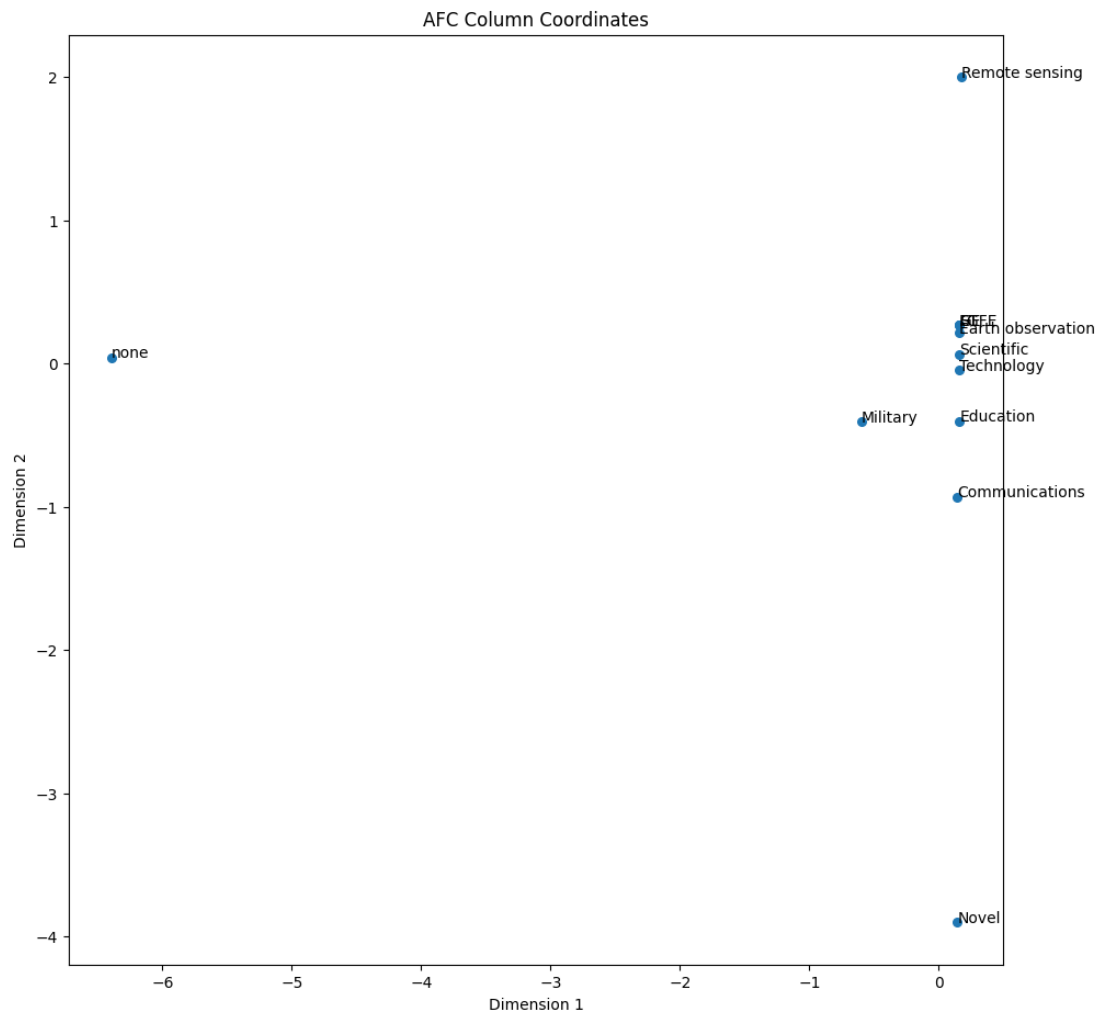


Figure 50 : Coordonnées de la colonne AFC



### 3.9 Cordonnée de la ligne et colonne AFC sans Australie

Afin de corriger le biais des données et des résultats on a amené de fixer cette erreur par l'élimination du pays Australie. Après avoir exclu ce pays on a remarqué que les pays sont très séparés et les pays qui sont proches ont presque le même type de catégorie des nanosatellites. En revanche les pays qui sont plus éloignés n'ont pas le même type de catégorie.

Cette exclusion nous permet de mieux comprendre les schémas et les relations entre les pays en nous concentrant sur ceux qui ont des caractéristiques plus similaires en matière de catégories de nanosatellites.

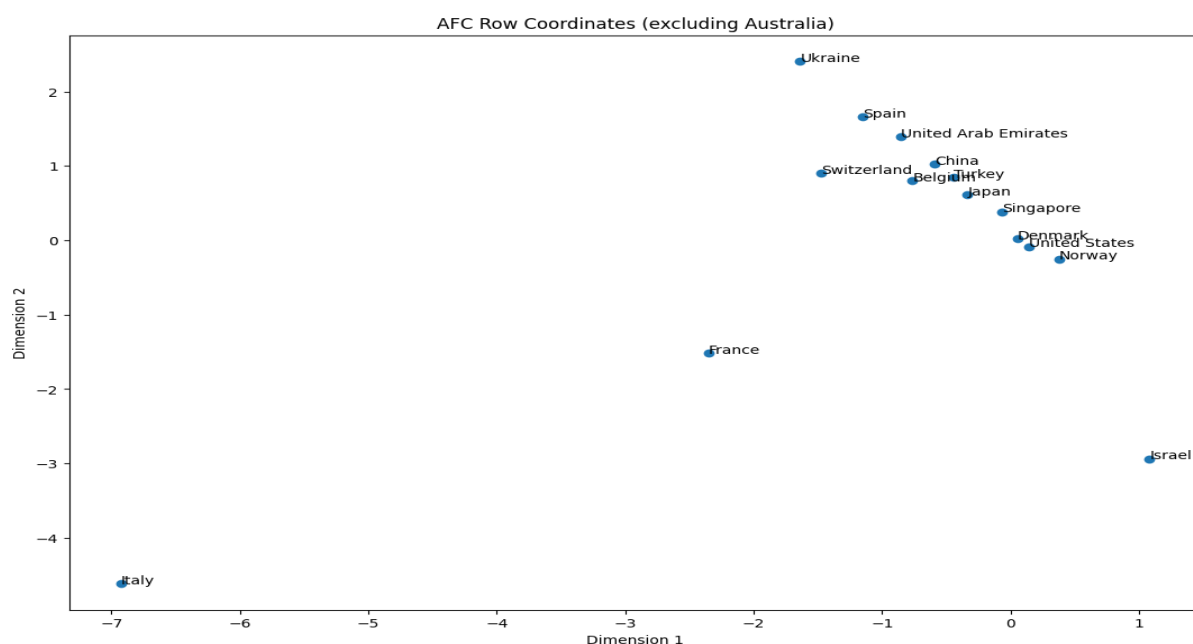


Figure 51 : Coordonnées de la ligne AFC sans Australie

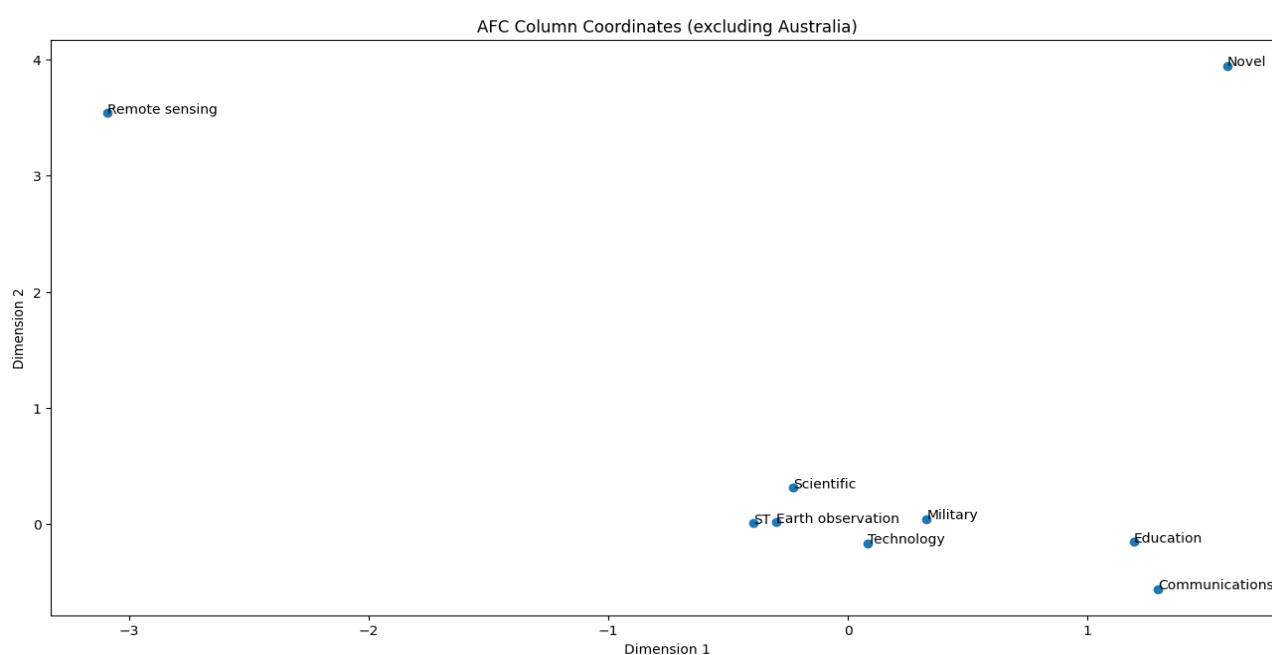


Figure 52 : Coordonnées de la colonne AFC sans Australie

# Conclusion

Au nom de ce modeste travail, nous tenons à souligner l'importance de ce stage de fin d'études pour notre développement professionnel. Il a été une étape irremplaçable dans notre parcours, nous permettant de façonner notre personnalité dans le domaine professionnel et de nous familiariser avec ce nouvel environnement rempli de surprises tant attendues qu'inattendues. Ce stage nous a également permis de prendre conscience des difficultés et des exigences propres à ce domaine.

Nous tenons particulièrement à souligner que notre expérience au sein de la société KSF Space a été inoubliable. Nous avons travaillé intensément, cherché des solutions et avons acquis une meilleure compréhension du travail en équipe. Nous avons réalisé des tâches individuelles tout en collaborant avec l'équipe en place au sein de l'entreprise. Nous avons eu la chance de travailler avec de vraies données liées aux nanosatellites en collaboration avec la NASA, ce qui était extrêmement captivant pour des étudiants en licence. Cette opportunité nous a été offerte sans aucun piston ou privilège, ce qui en fait une expérience encore plus précieuse.

Nous avons acquis de nouvelles notions, amélioré nos compétences en programmation, résolu des erreurs, et appris à analyser et interpréter les données. Nous sommes de retour plus vigilants et conscients que l'amélioration dans le domaine professionnel, que ce soit dans le développement web ou logiciel, nécessite beaucoup d'efforts et de patience. Nous tenons à souligner que l'auto-formation a été un élément clé dans ce parcours, renforçant nos compétences de manière significative.

Nous sommes reconnaissants de cette expérience enrichissante qui nous a permis d'élargir nos horizons professionnels et d'acquérir des compétences précieuses. Ce stage nous a ouvert les portes vers de nouvelles opportunités et nous a donné la confiance nécessaire pour poursuivre notre cheminement dans le domaine de la modélisation statistique informatique.

## WEBOGRAPHIE

<https://ksf.space/>

<https://www.nasa.gov/>

<https://www.math.univ-toulouse.fr/~besse/Wikistat/pdf/st-m-explo-afc.pdf>

<https://www.xlstat.com/fr/solutions/fonctionnalites/analyse-factorielle-des-correspondances-afc>

<https://www.openmindt.com/blog/dev/les-differents-types-de-tableaux-de-bord>

<https://www.ionos.fr/digitalguide/sites-internet/developpement-web/quest-ce-que-le-web-scraping/>

<https://www.nanosats.eu/database>