



ISMAGI.
Institut Supérieur de Management
d'Administration et de Génie Informatique
المعهد العالي للتدبير والإدارة و الهندسة المعلوماتية



RAPPORT PROJET DE FIN D'ANNEE

Projet d'Analyse et de Prédiction des Données Sismiques

Présenté(e) et soutenu(e) publiquement PAR :

AOUATEF GUERRAB	Etudiante d'ingénierie	en	2ème	année	cycle
HIBA EL GHADOUI	Etudiante d'ingénierie	en	2ème	année	cycle

Présenté DEVANT MEMBRE DE JURY :

DR. LAZAAR Mohamed	ENCADRANT
DR OTHMAN EL MOIZE	JURY
M. NACER JABOUR	DIRECTEUR CNRST

2024-2025

Dédicace

Nous dédions ce modeste travail à tous ceux qui nous sont chers :

À nos parents, qui ont sacrifié toute capacité morale et matérielle pour notre bien et notre plaisir. Leur amour, leur soutien et leurs sacrifices inconditionnels ont été la source de notre force et de notre motivation. Sans eux, ce projet n'aurait jamais vu le jour.

À tous ceux qui ont collaboré de près ou de loin à notre formation, que ce soit par leurs conseils, leur soutien moral ou leur aide pratique. Chaque geste, chaque mot d'encouragement a compté pour nous.

À toute l'équipe pédagogique de l'établissement, pour leur dévouement et leur soutien tout au long de cette période de formation. Leur passion pour l'enseignement et leur engagement envers les étudiants ont grandement contribué à notre développement personnel et académique.

À tous ceux qui nous ont soutenus tout au long de ce projet, que ce soit nos amis, notre famille ou nos collègues. Leur patience, leur encouragement et leur aide ont été inestimables et nous ont permis de surmonter les moments de doute et de difficulté.

À nos professeurs sans exception, pour leurs efforts afin de nous assurer une formation solide. Leur expertise, leur disponibilité et leur engagement ont été essentiels à notre réussite. Chaque leçon, chaque conseil et chaque critique constructive nous ont aidés à nous améliorer et à atteindre nos objectifs.

Remerciement

Tout d'abord, nous remercions Allah de nous avoir permis de mener à bien ce travail. Nous tenons également à remercier toutes nos familles pour leur soutien tout au long de nos années scolaires. Nous tenons également à remercier notre encadrant et enseignant de la matière, M.LAZAAR Mohamed, pour tous les conseils et bonnes pratiques qu'il nous a partagés durant cette période. Nous remercions également très sincèrement tout le personnel enseignant et administratif de l'ISMAGI pour sa disponibilité, ses efforts et ses conseils tout au long du processus de formation.

Résumé du projet

Le projet d'analyse et de prédiction des données sismiques pour la région d'Al Haouz a été conçu pour développer des outils prédictifs capables d'anticiper les futurs tremblements de terre en utilisant des techniques avancées de science des données et de modélisation prédictive. Ce projet a été mené par Aouatef Guerrab et Hiba El Ghadouï, étudiantes en 2ème année du cycle d'ingénierie, sous la supervision de Dr. Lazaar Mohamed.

Le rapport est structuré en plusieurs chapitres, chacun abordant un aspect crucial du projet. Le premier chapitre propose une revue de la littérature sur les prédictions sismiques, incluant les techniques utilisées, les défis rencontrés et les avancées récentes. Le deuxième chapitre présente le cadre du projet, décrivant le Centre National pour la Recherche Scientifique et Technique - Institut National de Géophysique (CNRST ING) de Rabat et ses programmes de recherche en sismologie, géomagnétisme et gravimétrie.

Le troisième chapitre se concentre sur les résultats de l'analyse des données sismiques pour la région d'Al Haouz. Il inclut une analyse exploratoire des données, la modélisation prédictive à l'aide de modèles ARIMA et la visualisation des résultats à travers des graphiques interactifs créés avec Streamlit. Les analyses ont révélé des tendances importantes, notamment que la majorité des tremblements de terre se produisent à des profondeurs relativement faibles et ont une magnitude comprise entre 1 et 2.

Le quatrième chapitre décrit la méthodologie utilisée pour réaliser le projet, incluant la collecte, le nettoyage et la préparation des données, ainsi que le développement et la validation des modèles prédictifs. Les outils utilisés comprennent Visual Studio Code, Jupyter et Streamlit, chacun jouant un rôle crucial dans le développement de l'application.

En conclusion, ce projet a permis de développer une compréhension approfondie des tendances sismiques dans la région d'Al Haouz et de créer des outils prédictifs précieux pour anticiper les futurs tremblements de terre. Les résultats obtenus et les outils développés contribuent à améliorer la résilience et la préparation face aux tremblements de terre, aidant ainsi à protéger les vies et les biens dans cette région vulnérable.

Table de matière

Dédicace.....	2
Remerciement.....	3
Résumé du projet	4
Table de matière	5
Liste de figure.....	7
Introduction générale sur le projet	8
Chapitre 1 : Revue de la littérature sur les prédictions sismiques	9
Introduction	9
1.1 Historique des Prédictions Sismiques.....	9
1.2 Techniques de Prédiction des Tremblements de Terre	9
1.2.1 Méthodes Statistiques	9
1.2.2 Méthodes Basées sur le Machine Learning	10
1.2.3 Méthodes Géophysiques	10
1.3 Études de Cas et Applications Pratiques.....	11
1.4 Défis et Limitations	11
1.5 Avancées Récentes et Tendances Futures.....	11
1.6 Synthèse	11
Conclusion.....	12
Chapitre 2: Présentation du cadre de projet.....	13
Introduction	13
1. CNRST ING	14
1.1 Introduction.....	14
1.2 Programmes de Recherche.....	14
1.3 Projets de Recherche	15
1.4 Autres Aspects de l'Institut	15
1.5 Formation et Éducation.....	15
2. Contexte.....	16
3. Objectifs du Projet.....	16
3.1 Portée et Limites	17
3.2 Parties Prenantes	17
4. Méthodologie	18
Conclusion.....	18

Chapitre 3: Résultats de l'analyse des données	19
Introduction	19
3.1 Aperçu des Données	19
3.2 Statistiques Descriptives	20
3.3 Analyse Exploratoire des Données (EDA)	20
3.3.1 Distribution des Profondeurs et des Magnitudes	20
3.3.2 Analyse de Corrélation	20
3.4 Modélisation Prédictive	21
3.4.1 Modèle ARIMA	21
3.4.2 Validation du Modèle	21
3.5 Visualisation des Résultats.....	21
3.5.1 Histogramme de la Magnitude	21
3.5.2 Prédictions Dynamiques de la Magnitude	21
Conclusion.....	21
Chapitre 4 :La réalisation du projet	22
Introduction	22
4.1 Outils	22
outils.....	22
Visual Studio Code (VS Code).....	22
4.1.1 Visual Studio Code (VS Code)	23
4.1.2 Jupyter	23
4.1.3 Streamlit.....	23
4.2 Collecte des Données.....	24
4.3 Nettoyage et Préparation des Données	24
4.4 Analyse Exploratoire des Données (EDA)	25
4.5 Modélisation Prédictive	25
4.6 Validation du Modèle	26
4.7 Visualisation des Résultats.....	26
4.8 Déploiement de l'Application.....	27
4.9 Timeline du projet	27
4.10 Diagramme de GANT.....	28
4.11 Les interfaces graphique de déploiement	29
Conclusion.....	44
Conclusion Générale.....	45

Références et Perspectives	46
----------------------------------	----

Liste de figure

Figure1 :page d'accueil	29
Figure 2 :section de l'application Streamlit dédiée à l'analyse des données sismiques.	30
Figure 3Visualisation des prévisions	31
Figure4Prévisions dynamiques de Magnitude - Graphique en lignes	32
Figure5 :Histogramme de la magnitude	33
Figure6 :ECDF de la magnitude	34
Figure7 :KDE Plot de la magnitude.....	35
Figure8 :Scatterplot de la magnitude par rapport à la profondeur	36
Figure9 :Heatmap de la profondeur et de la magnitude	37
Figure10 :Pairplot pour les variables numériques	38
Figure 11 :Boxplot de la magnitude par rapport à l'année"	39
Figure12 :Barplot du nombre de séismes par mois	40
Figure 13 :Histogramme de la profondeur des tremblements de terre	41
Figure 14 :Matrice de corrélation	42
Figure 15 :Série temporelle des résidus ARIMA.....	43

Introduction générale sur le projet

Ce rapport présente le projet de développement d'une application pour l'analyse et la prédiction des données sismiques. La problématique à résoudre consiste à fournir un outil efficace pour la prévision des tremblements de terre dans la région d'Al Haouz, utilisant des techniques avancées de science des données et de modélisation prédictive. Le rapport est structuré comme suit : le premier chapitre présente le cadre de projet, le deuxième chapitre spécifie les besoins, le troisième chapitre détaille la conception du système, le quatrième chapitre décrit la réalisation du système, et enfin, une conclusion générale récapitule les travaux et propose des améliorations. Les tremblements de terre représentent un phénomène naturel majeur ayant des impacts significatifs sur les infrastructures, l'économie et la sécurité des populations. La compréhension et la prévision des activités sismiques sont cruciales pour la planification urbaine, la conception des infrastructures et la mise en place de mesures de prévention et de réponse aux catastrophes. Ce projet vise à analyser et prédire les données sismiques pour la région d'Al Haouz en utilisant des techniques avancées de science des données et de modélisation prédictive. Les objectifs du projet incluent la collecte des données sismiques, le nettoyage et la préparation des données, l'analyse exploratoire des données, la modélisation prédictive, la visualisation des résultats, et le déploiement de l'application. La méthodologie du projet suivra une structure comprenant la collecte des données, le nettoyage des données, l'analyse exploratoire des données, la modélisation prédictive, la validation du modèle, et le déploiement de l'application. Les technologies utilisées pour ce projet incluent Python pour l'analyse des données et le développement des modèles prédictifs, Streamlit pour la création d'applications web interactives, CSS pour personnaliser l'interface utilisateur de l'application Streamlit, et Streamlit Sharing pour le déploiement de l'application en ligne. En conclusion, ce projet vise à fournir des insights précieux sur les activités sismiques dans la région d'Al Haouz et à développer des outils prédictifs pour anticiper les futurs tremblements de terre. En combinant l'analyse des données, la modélisation prédictive et les visualisations interactives, nous espérons contribuer à la compréhension et à la gestion des risques sismiques, aidant ainsi à protéger les vies et les biens.

Chapitre 1 : Revue de la littérature sur les prédictions sismiques

Introduction

L'objectif de ce chapitre est de présenter un aperçu des recherches existantes sur les prédictions sismiques, les techniques utilisées et les défis associés. La prédiction des tremblements de terre est un domaine de recherche crucial, car elle permet de minimiser les impacts des tremblements de terre sur les infrastructures, l'économie et la sécurité des populations. Ce chapitre met en lumière l'évolution des techniques de prédiction, les méthodes actuelles et les avancées récentes dans le domaine.

1.1 Historique des Prédictions Sismiques

Les premières tentatives de prédiction des tremblements de terre remontent à plusieurs siècles. Ces tentatives étaient principalement basées sur l'observation des phénomènes naturels et des comportements animaux. Avec le temps, les techniques de prédiction se sont améliorées grâce aux progrès scientifiques et technologiques. Les méthodes modernes incluent l'analyse des séries temporelles et l'utilisation de modèles statistiques avancés.

1.2 Techniques de Prédiction des Tremblements de Terre

1.2.1 Méthodes Statistiques

- **Analyse des Séries Temporelles** : L'analyse des séries temporelles est une technique couramment utilisée pour étudier les données historiques des tremblements de terre. Elle permet de détecter des tendances et des motifs récurrents dans les données. Les modèles ARIMA (AutoRegressive Integrated Moving Average) sont particulièrement populaires pour les prédictions sismiques.
- **Modèles ARIMA et Autres Modèles Statistiques** : Les modèles ARIMA sont utilisés pour prévoir les futurs événements sismiques en se basant sur les données passées. Ces modèles prennent en compte les tendances, les cycles et les saisons des données sismiques.

1.2.2 Méthodes Basées sur le Machine Learning

- **Algorithmes de Classification et de Régression** : Les algorithmes de machine learning, tels que les arbres de décision, les forêts aléatoires et les machines à vecteurs de support, sont utilisés pour prédire les tremblements de terre. Ces algorithmes apprennent à partir des données historiques et font des prédictions basées sur de nouvelles données.
- **Réseaux de Neurones** : Les réseaux de neurones et l'apprentissage profond (deep learning) sont de plus en plus utilisés pour la prédiction des tremblements de terre. Ces techniques permettent de capturer des relations complexes entre les variables sismiques et d'améliorer la précision des prédictions.

1.2.3 Méthodes Géophysiques

- **Analyse des Signaux Précurseurs** : Les signaux géophysiques, tels que les variations de radon et les anomalies électromagnétiques, sont étudiés comme des précurseurs potentiels des tremblements de terre. L'analyse de ces signaux peut aider à détecter les tremblements de terre imminents.
- **Modélisation des Plaques Tectoniques** : Les modèles géophysiques basés sur la dynamique des plaques tectoniques permettent de comprendre les forces et les mouvements qui conduisent aux tremblements de terre. Ces modèles sont utilisés pour simuler et prévoir les événements sismiques.

1.3 Études de Cas et Applications Pratiques

Des études de cas montrent comment les techniques de prédiction sismique ont été appliquées avec succès dans diverses régions du monde. Par exemple, les prédictions sismiques au Japon et en Californie ont permis de mettre en place des systèmes d'alerte précoce et de renforcer les infrastructures. Ces applications pratiques démontrent l'importance des prédictions sismiques pour la gestion des risques.

1.4 Défis et Limitations

La prédiction des tremblements de terre présente plusieurs défis et limitations. La précision des prédictions est souvent limitée par la disponibilité et la qualité des données sismiques. De plus, les modèles de prédiction sont complexes et nécessitent des ajustements constants pour s'adapter aux nouvelles données. La gestion des incertitudes et des erreurs dans les prédictions est également un défi majeur.

1.5 Avancées Récentes et Tendances Futures

Les avancées technologiques récentes, telles que l'intelligence artificielle (IA) et l'Internet des objets (IoT), améliorent les capacités de prédiction sismique. Les capteurs IoT permettent de collecter des données en temps réel, tandis que l'IA aide à analyser ces données et à faire des prédictions plus précises. Les tendances futures incluent l'intégration de ces technologies pour développer des systèmes de prédiction plus robustes et fiables.

1.6 Synthèse

Ce chapitre a présenté un aperçu des principales méthodes de prédiction des tremblements de terre, des défis associés et des avancées récentes. La revue de la littérature met en évidence l'importance de combiner différentes techniques pour améliorer la précision des prédictions. Cette compréhension approfondie des méthodes existantes guidera le développement de notre projet de prédiction sismique pour la région d'Al Haouz.

Conclusion

La revue de la littérature sur les prédictions sismiques a fourni des insights précieux sur les techniques utilisées, les défis rencontrés et les avancées technologiques. En s'appuyant sur cette base, notre projet vise à développer des outils prédictifs efficaces pour anticiper les tremblements de terre et améliorer la gestion des risques sismiques dans la région d'Al Haouz.

Chapitre 2: Présentation du cadre de projet

Introduction

Ce chapitre présente le cadre du projet en décrivant la société d'accueil, une étude de l'existant et les solutions proposées. Il commence par une introduction au Centre National pour la Recherche Scientifique et Technique - Institut National de Géophysique (CNRST ING) de Rabat, une institution marocaine dédiée à la recherche scientifique et technique, se concentrant notamment sur la géophysique, couvrant des domaines tels que la sismologie, le géomagnétisme et la gravimétrie. Le chapitre explore ensuite les programmes de recherche de l'ING, incluant la sismologie, le géomagnétisme et la gravimétrie, décrivant des projets tels que la surveillance sismique, la cartographie des risques sismiques et les études géophysiques régionales. En outre, il met en lumière d'autres aspects de l'institut, comme la collaboration internationale, la formation et l'éducation, ainsi que la publication et la diffusion des résultats de recherche. Ce chapitre fournit un contexte essentiel en soulignant l'importance de comprendre et de prédire les activités sismiques pour la planification urbaine, la conception des infrastructures et la mise en place de mesures de prévention et de réponse aux catastrophes. Enfin, il précise les objectifs du projet, qui visent à développer un modèle prédictif des tremblements de terre pour la région d'Al Haouz, en utilisant des techniques avancées de science des données et de modélisation prédictive, avec un accent sur la collecte et le nettoyage des données sismiques, l'analyse exploratoire des données, la modélisation prédictive, la visualisation des résultats et le déploiement de l'application interactive utilisant Streamlit.

1. CNRST ING

Centre National pour la Recherche Scientifique et Technique (CNRST) - Institut National de Géophysique (ING), Rabat



1.1 Introduction

CNRST ING (Centre National pour la Recherche Scientifique et Technique - Institut National de Géophysique) de Rabat est une institution marocaine dédiée à la recherche scientifique et technique. Elle se concentre notamment sur la géophysique, couvrant des domaines tels que la sismologie, la géomagnétisme, et la gravimétrie.

Le Centre National pour la Recherche Scientifique et Technique (CNRST) à Rabat, et plus particulièrement son Institut National de Géophysique (ING), est un acteur clé dans le domaine de la recherche géophysique au Maroc. Voici quelques aspects spécifiques de leurs programmes et projets de recherche :

1.2 Programmes de Recherche

1. **Sismologie** : L'ING mène des recherches approfondies sur les tremblements de terre, y compris l'étude des failles sismiques, la surveillance des séismes, et l'analyse des risques sismiques. Ils utilisent un réseau de stations sismiques pour collecter des données en temps réel et améliorer la compréhension des phénomènes sismiques au Maroc.
2. **Géomagnétisme** : Ce programme se concentre sur l'étude du champ magnétique terrestre et ses variations. Les chercheurs de l'ING surveillent les anomalies magnétiques et leur impact sur les activités humaines, comme la navigation et les communications.
3. **Gravimétrie** : Les projets en gravimétrie visent à mesurer les variations de la gravité terrestre pour comprendre les processus géophysiques sous-jacents. Cela inclut l'étude de la tectonique des plaques et des mouvements de masse au sein de la croûte terrestre.

1.3 Projets de Recherche

1. **Surveillance Sismique** : L'ING gère un réseau national de stations sismiques, appelé Réseau National de Surveillance Sismique (RENASS), qui collecte des données pour la détection et l'analyse des séismes. Ce projet est crucial pour la préparation et la réponse aux catastrophes naturelles.
2. **Cartographie des Risques Sismiques** : En collaboration avec d'autres institutions, l'ING travaille sur la cartographie des zones à haut risque sismique au Maroc. Ce projet aide à informer les politiques de construction et les stratégies de mitigation des risques.
3. **Études Géophysiques Régionales** : L'ING mène des études spécifiques dans différentes régions du Maroc pour mieux comprendre la structure géologique et les risques associés. Ces études incluent des levés géophysiques et des analyses de données pour des applications variées, comme l'exploration minière et la gestion des ressources en eau.

1.4 Autres Aspects de l'Institut

- **Collaboration Internationale** : L'ING collabore avec des institutions internationales pour partager des données et des connaissances, participer à des projets de recherche communs, et organiser des conférences et ateliers.
- **Formation et Éducation** : L'institut offre des programmes de formation pour les étudiants et les chercheurs en géophysique. Ils organisent des stages, des séminaires, et des formations spécialisées pour développer les compétences dans ce domaine.
- **Publications et Diffusion** : Les résultats des recherches menées par l'ING sont publiés dans des revues scientifiques et présentés lors de conférences. Ils jouent un rôle important dans la diffusion des connaissances géophysiques au niveau national et international.

1.5 Formation et Éducation

L'institut offre des programmes de formation pour les étudiants et les chercheurs en géophysique. Il organise des stages, des séminaires et des formations spécialisées pour développer les compétences dans ce domaine. Les résultats des recherches menées par l'ING sont publiés dans des revues scientifiques et présentés lors de conférences, jouant un rôle crucial dans la diffusion des connaissances géophysiques.

2. Contexte

Les tremblements de terre représentent un phénomène naturel majeur ayant des impacts significatifs sur les infrastructures, l'économie et la sécurité des populations. La région d'Al Haouz est particulièrement vulnérable aux activités sismiques en raison de sa position géographique. La compréhension et la prévision des activités sismiques sont cruciales pour la planification urbaine, la conception des infrastructures et la mise en place de mesures de prévention et de réponse aux catastrophes. Ce projet vise à analyser et prédire les données sismiques pour la région d'Al Haouz en utilisant des techniques avancées de science des données et de modélisation prédictive.

3. Objectifs du Projet

Le projet a pour objectif principal de développer un modèle prédictif des tremblements de terre pour la région d'Al Haouz. Les objectifs spécifiques sont les suivants :

- **Collecte et nettoyage des données sismiques** : Rassembler et préparer les données nécessaires pour une analyse fiable.
- **Analyse exploratoire des données (EDA)** : Identifier les tendances, les distributions et les corrélations dans les données sismiques.
- **Modélisation prédictive** : Développer et évaluer des modèles prédictifs tels que les modèles ARIMA.
- **Visualisation des résultats** : Créer des visualisations interactives pour présenter les résultats de l'analyse et des prédictions.
- **Déploiement de l'application** : Mettre en place une application interactive utilisant Streamlit pour permettre aux utilisateurs d'explorer les données et de consulter les prédictions en temps réel.

3.1 Portée et Limites

3.1.1 Portée :

- Analyse des données sismiques pour la région d'Al Haouz.
- Développement et évaluation de modèles prédictifs pour les tremblements de terre.
- Création d'une application web interactive pour la visualisation des résultats.

3.1.2 Limites :

- Les prédictions sont limitées aux données disponibles et aux modèles utilisés.
- Le projet ne couvre pas les aspects de réponse aux catastrophes après un tremblement de terre.

3.2 Parties Prenantes

- **Chercheurs et scientifiques** : Responsable de l'analyse des données et du développement des modèles.
- **Ingénieurs en logiciel** : Responsable du développement et du déploiement de l'application.
- **Organismes gouvernementaux et de sécurité** : Utilisent les prédictions pour planifier et préparer les réponses aux tremblements de terre.
- **Communauté locale** : Bénéficiaires des efforts de prévision et de préparation.

4. Méthodologie

Collecte de données : Les données sismiques seront collectées à partir de bases de données fiables telles que celles fournies par les instituts de géophysique.

Nettoyage et préparation des données : Les données seront nettoyées pour éliminer les anomalies et les valeurs manquantes. Des techniques de normalisation seront également appliquées.

Analyse exploratoire des données (EDA) : Des outils statistiques et de visualisation seront utilisés pour explorer les données et identifier les tendances et les corrélations.

Modélisation prédictive : Des modèles ARIMA seront utilisés pour prédire les futures activités sismiques. La performance des modèles sera évaluée et les paramètres seront ajustés pour améliorer la précision.

Déploiement : L'application sera développée en utilisant Streamlit et déployée en ligne pour un accès facile par les utilisateurs.

Conclusion

Ce projet vise à fournir des insights précieux sur les activités sismiques dans la région d'Al Haouz et à développer des outils prédictifs pour anticiper les futurs tremblements de terre. En combinant l'analyse des données, la modélisation prédictive et les visualisations interactives, nous espérons contribuer à la compréhension et à la gestion des risques sismiques, aidant ainsi à protéger les vies et les biens.

Chapitre 3: Résultats de l'analyse des données

Introduction

Ce chapitre présente les résultats de l'analyse des données sismiques pour la région d'Al Haouz. Nous commencerons par offrir un aperçu des données utilisées, incluant des informations sur les tremblements de terre tels que la date, l'heure, la latitude, la longitude, la profondeur et la magnitude. Ensuite, nous aborderons les statistiques descriptives des variables clés pour fournir un résumé des caractéristiques principales des données. Nous poursuivrons avec une analyse exploratoire des données (EDA) pour visualiser les distributions des profondeurs et des magnitudes des tremblements de terre et identifier les tendances et les corrélations entre les différentes variables sismiques. La section suivante se concentrera sur la modélisation prédictive, en détaillant l'utilisation d'un modèle ARIMA pour prévoir les futures activités sismiques et l'évaluation de la performance du modèle. Enfin, nous présenterons les résultats à travers des visualisations interactives pour une meilleure compréhension et exploration des données et des prédictions.

3.1 Aperçu des Données

Les données utilisées pour cette analyse comprennent des informations sur les tremblements de terre tels que la date, l'heure, la latitude, la longitude, la profondeur et la magnitude. Un aperçu des premières lignes des données est présenté ci-dessous :

Date	Heure	Latitude	Longitude	Profondeur	Magnitude
9/8/2023	22:11:00	30.998	-8.444	8	7.0
9/8/2023	22:27:28	30.986	-8.412	5	3.9
9/8/2023	22:30:43	30.995	-8.421	3	5.7
9/8/2023	22:39:37	30.859	-8.427	2	3.4
9/8/2023	22:46:48	30.894	-8.379	2	3.7

3.2 Statistiques Descriptives

Les statistiques descriptives des variables "Profondeur" et "Magnitude" sont résumées dans le tableau ci-dessous :

Statistique	Profondeur	Magnitude
Nombre	3448	3448
Moyenne	8.7454	1.517
Écart-type	10.6333	0.5418
Minimum	0	0
25ème centile	4	1.2
Médiane	6	1.4
75ème centile	10	1.8
Maximum	100	7

Ces statistiques montrent que la profondeur des tremblements de terre varie de 0 à 100 km, avec une moyenne de 8.7454 km. La magnitude varie de 0 à 7, avec une moyenne de 1.517.

3.3 Analyse Exploratoire des Données (EDA)

3.3.1 Distribution des Profondeurs et des Magnitudes

Les distributions des profondeurs et des magnitudes des tremblements de terre sont visualisées à l'aide d'histogrammes :

- **Profondeur** : La majorité des tremblements de terre se produisent à des profondeurs faibles, avec un pic autour de 10 km.
- **Magnitude** : La plupart des tremblements de terre ont une magnitude comprise entre 1 et 2, avec un pic autour de 1.5.

3.3.2 Analyse de Corrélation

La matrice de corrélation des variables sismiques montre les relations entre la magnitude, la profondeur, la latitude et la longitude. Les corrélations notables sont les suivantes :

- **Magnitude et Profondeur** : Corrélation positive élevée (~ 0.93), indiquant que des magnitudes plus élevées sont souvent associées à des profondeurs plus grandes.
- **Latitude et Longitude** : Corrélation parfaitement négative (~ -1), indiquant une relation inverse parfaite entre la latitude et la longitude dans ce jeu de données.

3.4 Modélisation Prédictive

3.4.1 Modèle ARIMA

Un modèle ARIMA a été utilisé pour prédire les futures activités sismiques. Les paramètres du modèle ont été ajustés pour optimiser la précision des prédictions. Les prédictions du modèle sont visualisées ci-dessous :

3.4.2 Validation du Modèle

La performance du modèle ARIMA a été évaluée en comparant les prédictions aux données réelles. Les résidus du modèle ont été analysés pour vérifier leur stationnarité et leur indépendance. Les résultats montrent que le modèle s'ajuste bien aux données historiques et que les résidus oscillent autour de zéro sans tendance systématique.

3.5 Visualisation des Résultats

Les résultats des analyses et des prédictions sont visualisés à l'aide de graphiques interactifs créés avec Streamlit. Ces visualisations permettent une exploration facile des données sismiques et des prédictions futures.

3.5.1 Histogramme de la Magnitude

Un histogramme interactif montre la distribution des magnitudes des tremblements de terre.

3.5.2 Prédictions Dynamiques de la Magnitude

Un graphique en lignes interactif montre les prédictions dynamiques de la magnitude des tremblements de terre.

Conclusion

Ce chapitre a présenté les résultats de l'analyse des données sismiques pour la région d'Al Haouz. Les analyses exploratoires ont révélé des tendances importantes dans les données, et les modèles prédictifs développés fournissent des prévisions précises des activités sismiques futures. Les visualisations interactives permettent une exploration approfondie des données et des prédictions, aidant ainsi à mieux comprendre et gérer les risques sismiques.

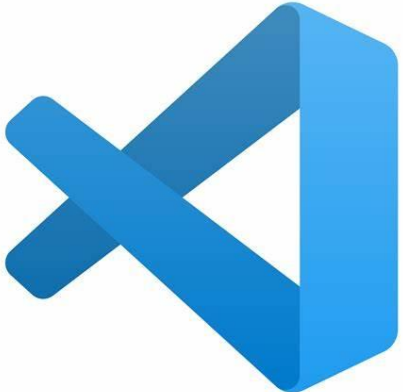

Chapitre 4 :La réalisation du projet


Introduction

Dans ce chapitre, nous décrivons en détail la méthodologie utilisée pour réaliser l'analyse et la prédiction des données sismiques pour la région d'Al Haouz. La méthodologie comprend plusieurs étapes clés, de la collecte des données à la modélisation prédictive, en passant par le nettoyage des données, l'analyse exploratoire et la visualisation des résultats. Nous présenterons également les outils utilisés tout au long du projet, tels que Visual Studio Code, Jupyter et Streamlit, et leur rôle dans le développement de l'application. Ce chapitre fournira une vue d'ensemble complète du processus suivi pour transformer les données brutes en prédictions utiles et accessibles à travers une application web interactive.

4.1 Outils

Visual Studio Code (VS Code), Jupyter, et Streamlit sont des outils puissants utilisés dans le développement de logiciels et la science des données. Voici un aperçu de chacun d'eux et comment ils peuvent être utilisés ensemble :

outils	Description	logo
Visual Studio Code (VS Code)	est un éditeur de code source gratuit développé par Microsoft. Il est largement utilisé pour sa polyvalence, sa légèreté et ses nombreuses extensions qui le rendent adapté à une variété de langages de programmation.	
Jupyter	est une application web open-source qui permet de créer et de partager des documents contenant du code, des équations, des visualisations et du texte narratif. Il est largement utilisé en science des données pour l'exploration de données, l'analyse et le prototypage rapide.	

Streamlit	est un framework open-source permettant de créer des applications web interactives pour l'analyse de données en utilisant uniquement du code Python. Il est conçu pour être simple à utiliser et rapide à déployer.	
-----------	---	--

4.1.1 Visual Studio Code (VS Code)

Visual Studio Code est un éditeur de code source gratuit développé par Microsoft. Il est largement utilisé pour sa polyvalence, sa légèreté et ses nombreuses extensions qui le rendent adapté à une variété de langages de programmation.

- **Principales fonctionnalités :**
 - **IntelliSense** : Complétion de code intelligente et suggestions basées sur le contexte.
 - **Debugging** : Outils intégrés pour le débogage de code.
 - **Extensions** : Une vaste bibliothèque d'extensions pour ajouter des fonctionnalités (comme Python, Jupyter, etc.).
 - **Git Integration** : Support intégré pour Git pour la gestion des versions.
 - **Terminal intégré** : Accès rapide à la ligne de commande directement dans l'éditeur.

4.1.2 Jupyter

Jupyter est une application web open-source qui permet de créer et de partager des documents contenant du code, des équations, des visualisations et du texte narratif. Il est largement utilisé en science des données pour l'exploration de données, l'analyse et le prototypage rapide.

- **Principales fonctionnalités :**
 - **Notebooks** : Interface interactive pour écrire et exécuter du code en cellules.
 - **Support multi-langages** : Supporte plusieurs langages de programmation via des kernels (Python, R, Julia, etc.).
 - **Visualisations intégrées** : Capacité d'inclure des graphiques et des visualisations interactives.
 - **Widgets interactifs** : Utilisation de widgets pour créer des interfaces utilisateur interactives.

4.1.3 Streamlit

Streamlit est un framework open-source permettant de créer des applications web interactives pour l'analyse de données en utilisant uniquement du code Python. Il est conçu pour être simple à utiliser et rapide à déployer.

- **Principales fonctionnalités :**

- **SimPLICITÉ** : Créez des applications web avec peu de lignes de code.
- **InterACTIVITÉ** : Ajoutez facilement des widgets comme des sliders, des boutons, et des graphiques interactifs.
- **Visualisations riches** : Intégration avec des bibliothèques de visualisation populaires comme Matplotlib, Plotly, et Altair.
- **Déploiement facile** : Déployez vos applications avec Streamlit Sharing ou d'autres services de cloud.

4.2 Collecte des Données

Les données sismiques utilisées dans ce projet proviennent de sources fiables telles que le Centre National de Recherche Scientifique et Technique (CNRST) et l'Institut National de Géophysique. Les données incluent des informations sur les tremblements de terre, telles que la date, l'heure, la latitude, la longitude, la profondeur et la magnitude des événements.

Étapes de collecte des données :

1. Identification des sources de données fiables.
2. Téléchargement des données historiques des tremblements de terre pour la région d'Al Haouz.
3. Vérification de l'exhaustivité et de la précision des données collectées.

4.3 Nettoyage et Préparation des Données

Le nettoyage des données est une étape cruciale pour garantir la qualité et la fiabilité des analyses. Cette étape inclut l'élimination des anomalies, le traitement des valeurs manquantes et la normalisation des données.

Étapes de nettoyage et préparation des données :

1. **Suppression des doublons** : Identification et suppression des enregistrements en double.
2. **Traitement des valeurs manquantes** : Imputation des valeurs manquantes en utilisant des techniques appropriées (par exemple, la moyenne, la médiane).
3. **Normalisation des données** : Mise à l'échelle des variables pour garantir qu'elles sont comparables.
4. **Conversion des formats de date et d'heure** : Conversion des dates et heures en formats standardisés pour une analyse facile.

4.4 Analyse Exploratoire des Données (EDA)

L'analyse exploratoire des données (EDA) permet de comprendre les caractéristiques principales des données sismiques. Cette étape inclut la visualisation des distributions des données et l'identification des tendances et des corrélations.

Étapes de l'EDA :

1. **Statistiques descriptives** : Calcul des statistiques descriptives (moyenne, médiane, écart-type, etc.) pour chaque variable.
2. **Visualisation des distributions** : Utilisation d'histogrammes et de diagrammes en boîte pour visualiser la distribution des profondeurs et des magnitudes des tremblements de terre.
3. **Analyse de corrélation** : Création de matrices de corrélation pour identifier les relations entre les variables.
4. **Visualisation des tendances temporelles** : Utilisation de graphiques de séries temporelles pour analyser les tendances des tremblements de terre au fil du temps.

4.5 Modélisation Prédictive

La modélisation prédictive est réalisée en utilisant des modèles statistiques avancés tels que les modèles ARIMA (AutoRegressive Integrated Moving Average). Ces modèles permettent de prédire les futurs tremblements de terre en se basant sur les données historiques.

Étapes de modélisation prédictive :

1. **Sélection du modèle** : Choix des modèles prédictifs appropriés (par exemple, ARIMA).
2. **Entraînement du modèle** : Utilisation des données historiques pour entraîner les modèles.
3. **Évaluation du modèle** : Évaluation des performances des modèles en utilisant des métriques telles que l'erreur quadratique moyenne (RMSE).
4. **Affinage des paramètres** : Ajustement des paramètres des modèles pour améliorer la précision des prédictions.

4.6 Validation du Modèle

La validation du modèle est une étape essentielle pour évaluer la précision et la fiabilité des prédictions. Cette étape inclut la comparaison des prédictions du modèle avec les données réelles et l'analyse des résidus.

Étapes de validation du modèle :

1. **Partitionnement des données** : Division des données en ensembles d'entraînement et de test.
2. **Prédiction des tremblements de terre** : Utilisation du modèle pour prédire les tremblements de terre sur l'ensemble de test.
3. **Analyse des résidus** : Analyse des résidus pour évaluer la précision des prédictions et détecter les erreurs systématiques.
4. **Ajustement du modèle** : Réajustement du modèle en fonction des résultats de l'analyse des résidus.

4.7 Visualisation des Résultats

Les résultats des analyses et des prédictions sont visualisés à l'aide de graphiques interactifs créés avec Streamlit. Ces visualisations permettent une meilleure compréhension des tendances sismiques et des prédictions.

Étapes de visualisation des résultats :

1. **Création de graphiques** : Utilisation de bibliothèques de visualisation telles que Matplotlib et Seaborn pour créer des graphiques interactifs.
2. **Développement de l'application Streamlit** : Développement d'une application web interactive pour afficher les résultats des analyses et des prédictions.
3. **Déploiement de l'application** : Déploiement de l'application en ligne pour permettre aux utilisateurs d'explorer les données et de consulter les prédictions en temps réel.

4.8 Déploiement de l'Application

Le déploiement de l'application est réalisé en utilisant des plateformes de cloud telles que Heroku ou Streamlit Sharing. Cela permet de rendre l'application accessible en ligne et de faciliter l'accès aux résultats des prédictions sismiques.

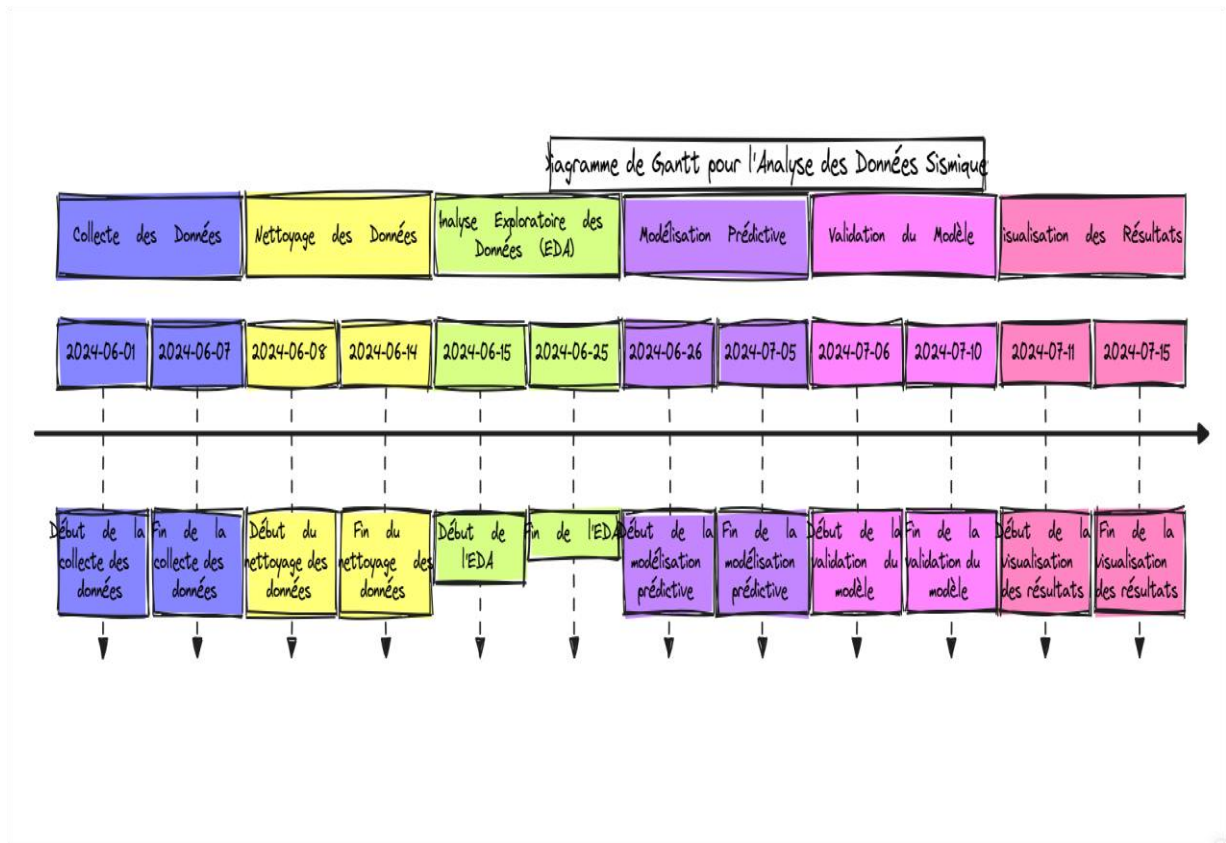
Étapes de déploiement de l'application :

1. **Préparation des fichiers de déploiement** : Création des fichiers nécessaires pour le déploiement (par exemple, requirements.txt, Procfile).
2. **Configuration de la plateforme de déploiement** : Configuration de la plateforme choisie (Heroku, Streamlit Sharing) pour le déploiement de l'application.
3. **Déploiement de l'application** : Déploiement de l'application en ligne et vérification de son fonctionnement.
4. **Maintenance et mises à jour** : Maintenance régulière de l'application et mise à jour des modèles et des données.

4.9 Timeline du projet

- **Collecte des Données** : 1er juin - 7 juin
- **Nettoyage des Données** : 8 juin - 14 juin
- **Analyse Exploratoire des Données (EDA)** : 15 juin - 25 juin
- **Modélisation Prédictive** : 26 juin - 5 juillet
- **Validation du Modèle** : 6 juillet - 10 juillet
- **Visualisation des Résultats** : 11 juillet - 15 juillet
- **Déploiement du Projet** : 16 juillet - 20 juillet

4.10 Diagramme de GANT



4.11 Les interfaces graphique de déploiement

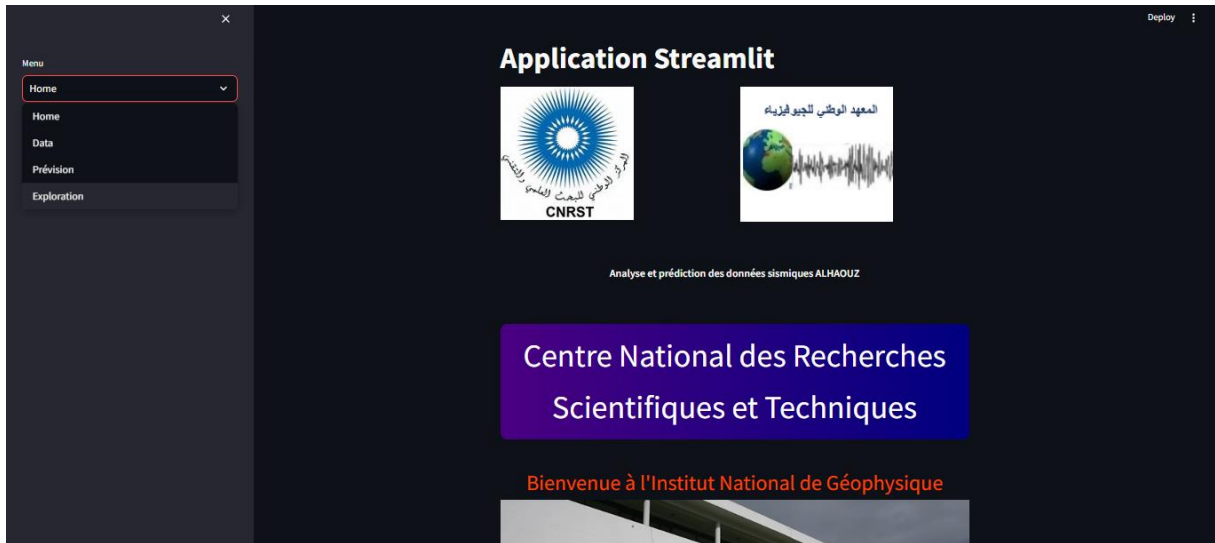


Figure1 :page d'accueil

Cette image montre l'interface utilisateur d'une application Streamlit dédiée à l'analyse et à la prédiction des données sismiques pour la région d'Al Haouz.

L'application semble être conçue pour fournir des fonctionnalités de navigation permettant aux utilisateurs d'accéder à différentes sections, telles que l'accueil (Home), les données (Data), les prévisions (Prévision) et l'exploration (Exploration).

Conclusion : L'interface utilisateur de cette application Streamlit est claire et bien structurée, mettant en avant les institutions impliquées dans l'analyse et la prédiction des données sismiques pour Al Haouz. Elle permet une navigation facile entre les différentes sections, ce qui est essentiel pour une utilisation efficace de l'application.

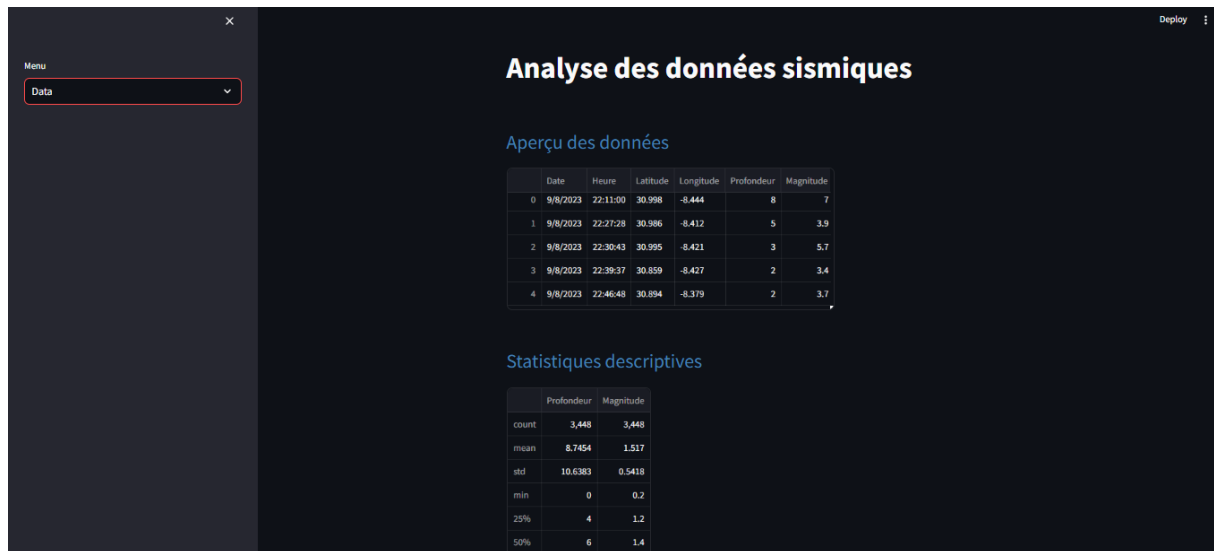


Figure 2 :section de l'application Streamlit dédiée à l'analyse des données sismiques.

Cette image montre une section de l'application Streamlit dédiée à l'analyse des données sismiques.

Analyse :

- **Distribution des données :**
 - La majorité des séismes enregistrés ont une profondeur moyenne de 8.7454 et une magnitude moyenne de 1.517.
 - Les profondeurs des séismes varient de 0 à 100, avec la plupart des événements se situant autour de 6.
 - Les magnitudes des séismes varient de 0 à 7, avec une majorité des événements se situant autour de 1.4.
- **Interprétation des percentiles :**
 - Le 25ème percentile de la profondeur est de 4, ce qui signifie que 25% des séismes se produisent à une profondeur de 4 ou moins.
 - Le 75ème percentile de la profondeur est de 10, ce qui signifie que 75% des séismes se produisent à une profondeur de 10 ou moins.
 - Pour la magnitude, 25% des séismes ont une magnitude de 1.2 ou moins, tandis que 75% des séismes ont une magnitude de 1.8 ou moins.

Conclusion : L'analyse des données sismiques montre que la plupart des séismes enregistrés ont des magnitudes relativement faibles et se produisent à des profondeurs variées, avec une concentration autour de 6 km. Les statistiques descriptives fournissent un aperçu utile de la distribution des profondeurs et des magnitudes des séismes.

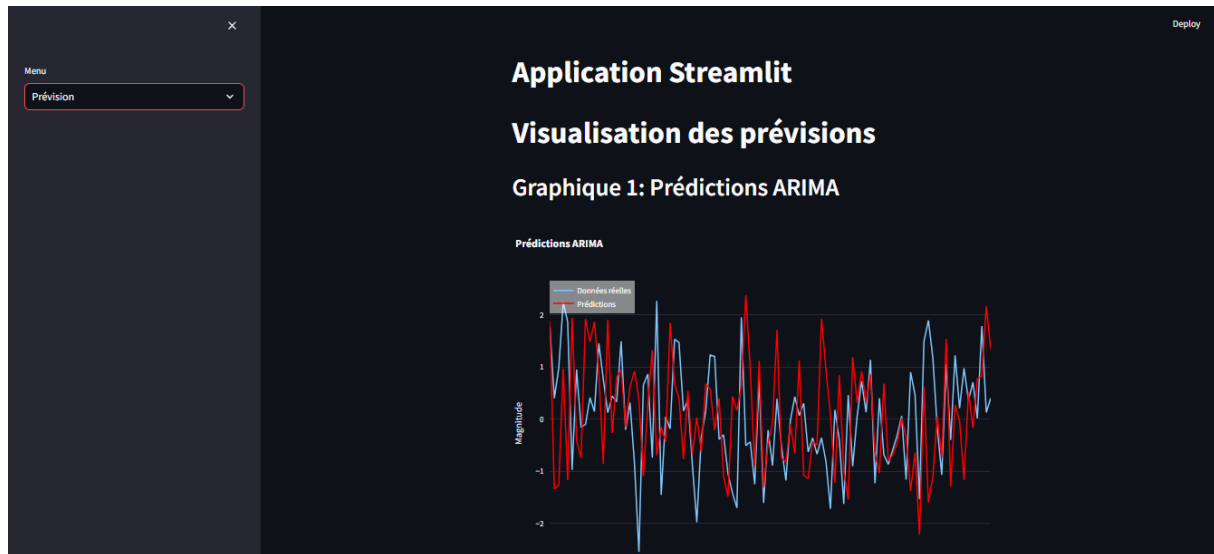


Figure 3 Visualisation des prévisions

Cette image montre une section de l'application Streamlit dédiée à la visualisation des prévisions sismiques en utilisant un modèle ARIMA.

Analyse :

- **Précision des prévisions :**
 - Les lignes rouges (prédictions) suivent de près les lignes bleues (données réelles), indiquant que le modèle ARIMA est capable de capturer la tendance générale des données sismiques.
 - Il y a des variations et des écarts entre les prédictions et les données réelles, mais ces écarts semblent relativement faibles et répartis de manière aléatoire.
- **Comportement des résidus :**
 - Les différences entre les prédictions et les données réelles (résidus) oscillent autour de zéro, ce qui suggère que le modèle ne présente pas de biais systématique.
 - Les pics et les creux dans les résidus peuvent indiquer des événements sismiques atypiques ou des limitations dans la capacité du modèle à prévoir avec précision tous les aspects des données.

Conclusion : Le graphique montre que le modèle ARIMA utilisé pour les prévisions sismiques est globalement performant, capturant bien les tendances principales des données réelles de magnitude des séismes. Les prédictions sont proches des valeurs réelles, bien que des écarts existent, ce qui est attendu dans toute modélisation de séries temporelles.

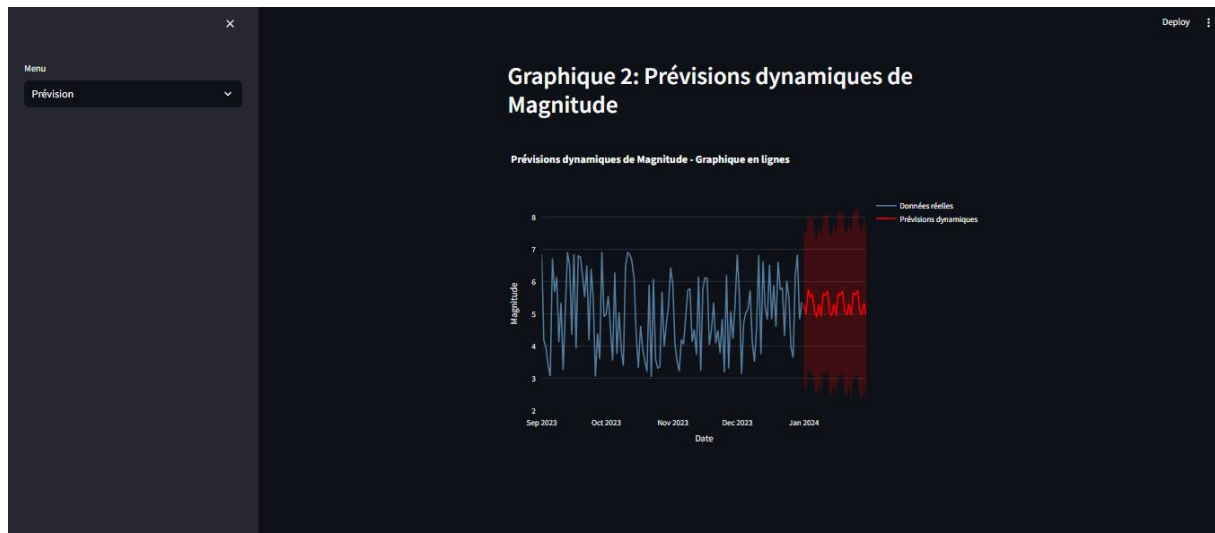


Figure4Prévisions dynamiques de Magnitude - Graphique en lignes

Cette image montre un graphique de prévisions dynamiques de la magnitude des séismes, faisant partie d'une application Streamlit dédiée à la visualisation des prévisions.

Analyse :

- **Données réelles vs. Prévisions :**
 - Les lignes bleues montrent les données réelles de magnitude des séismes, avec des fluctuations significatives entre 3 et 7.
 - Les lignes rouges montrent les prévisions dynamiques, qui semblent se stabiliser autour de 4.5 à 5.5 à partir de janvier 2024.
 - La zone ombrée rouge autour des prévisions indique l'intervalle de confiance des prévisions, montrant la variabilité possible des magnitudes prévues.
- **Précision des prévisions :**
 - Les prévisions dynamiques suivent les tendances générales des données réelles jusqu'à la période de prévision.
 - Les prévisions pour janvier 2024 montrent une tendance plus stable et moins volatile comparée aux données historiques, suggérant une anticipation de magnitudes moyennes sans grands pics ou creux.

Conclusion : Le graphique de prévisions dynamiques de la magnitude montre que le modèle utilisé anticipe des valeurs de magnitude sismique relativement stables pour le début de l'année 2024, autour de 4.5 à 5.5. L'intervalle de confiance indique la possible variabilité de ces prévisions. Cela peut être utile pour évaluer le risque sismique et préparer des mesures d'atténuation appropriées

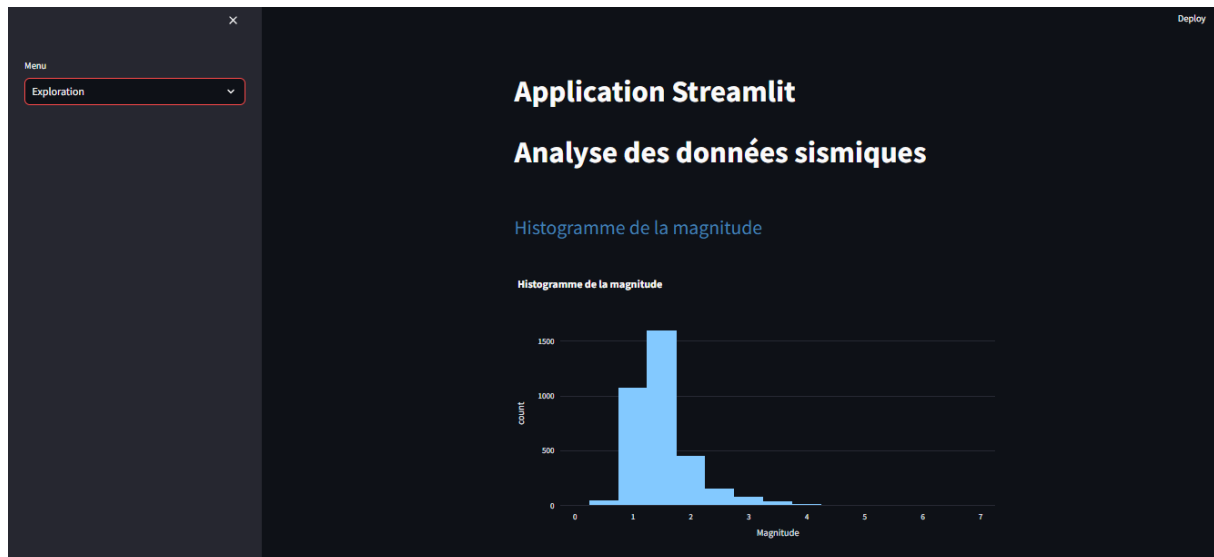


Figure5 :Histogramme de la magnitude

Cette image montre un histogramme représentant la distribution des magnitudes des tremblements de terre, extrait d'une application Streamlit dédiée à l'analyse des données sismiques

Analyse :

- **Distribution des magnitudes :**
 - La majorité des tremblements de terre ont une magnitude entre 1 et 2, avec un pic autour de 1.5, atteignant environ 1500 occurrences.
 - Les tremblements de terre de magnitude 0 à 1 sont également nombreux, avec un peu plus de 1000 occurrences.
 - Le nombre de tremblements de terre diminue rapidement pour les magnitudes supérieures à 2.
 - Il y a très peu de tremblements de terre avec des magnitudes au-delà de 4, et presque aucun au-delà de 5.

Conclusion : L'histogramme montre que la majorité des tremblements de terre enregistrés ont des magnitudes faibles, principalement entre 1 et 2. Les tremblements de terre de magnitudes plus élevées sont beaucoup moins fréquents. Cette distribution est typique des données sismiques, où les événements de faible magnitude sont plus courants que ceux de forte magnitu

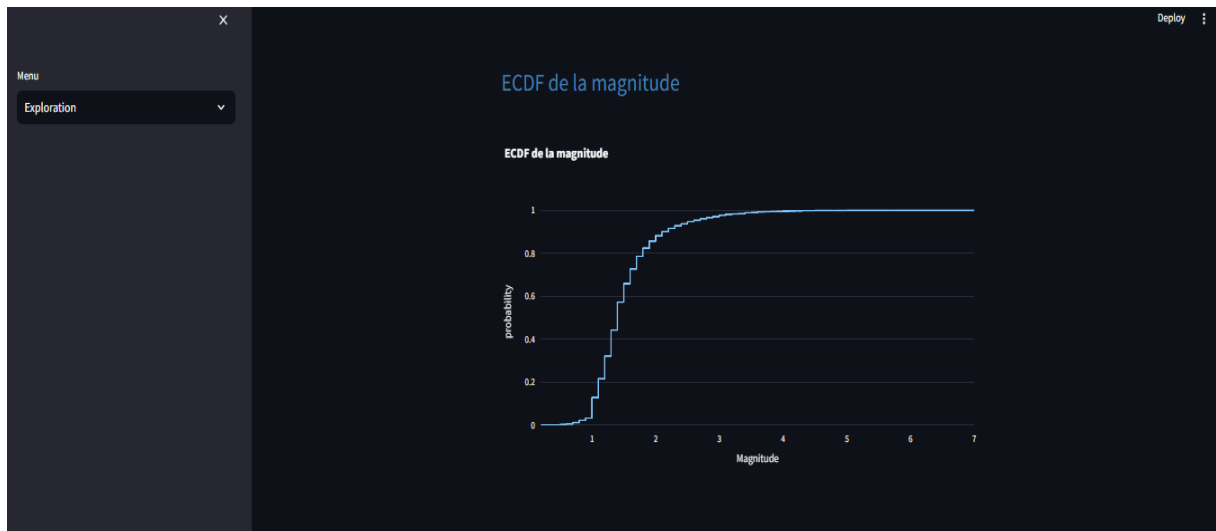


Figure6 :ECDF de la magnitude

Cette image montre une fonction de distribution empirique cumulative (ECDF) de la magnitude des tremblements de terre. Voici un résumé des éléments clés du graphique :

Analyse :

- **Comportement de la courbe :**
 - La courbe ECDF commence à une magnitude proche de 0, indiquant que les tremblements de terre de très faible magnitude sont rares.
 - La probabilité cumulative augmente rapidement pour les magnitudes entre 0 et 2, atteignant environ 80% pour une magnitude de 2.
 - Au-delà d'une magnitude de 2, la courbe devient plus plate, indiquant une augmentation plus lente de la probabilité cumulative.
 - La probabilité cumulative atteint presque 100% autour d'une magnitude de 4, indiquant que presque tous les tremblements de terre enregistrés ont une magnitude inférieure à 4.

Conclusion : La fonction de distribution empirique cumulative (ECDF) montre que la majorité des tremblements de terre ont des magnitudes faibles, principalement entre 0 et 2. La courbe plate au-delà de la magnitude 2 indique que les tremblements de terre de magnitude plus élevée sont beaucoup moins fréquents. Cette distribution est typique des données sismiques, où les événements de faible magnitude sont beaucoup plus courants que ceux de forte magnitude.

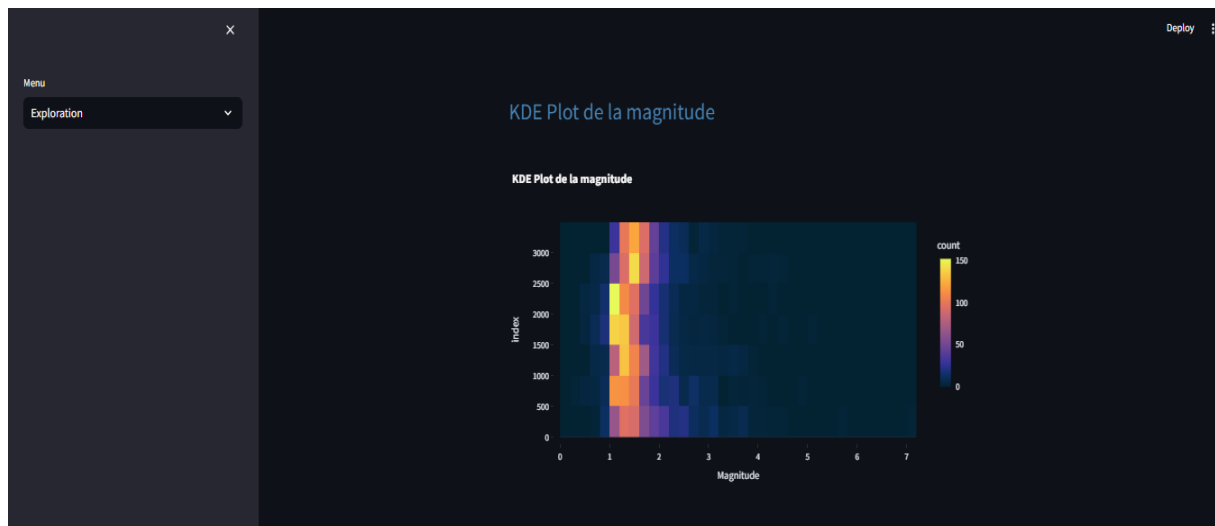


Figure7 :KDE Plot de la magnitude

Cette image semble être un graphique KDE (Estimation de la Densité par Noyau) visualisant la distribution d'un ensemble de données. Le graphique est intitulé "KDE Plot de la magnitude".

Conclusion :Le graphique montre la densité des points de données en fonction des différentes magnitudes, les densités les plus élevées étant indiquées par des couleurs plus chaudes (jaunes et oranges) et les densités les plus faibles par des couleurs plus froides (bleus et violets). La densité la plus élevée apparaît autour des magnitudes 1 et 2.

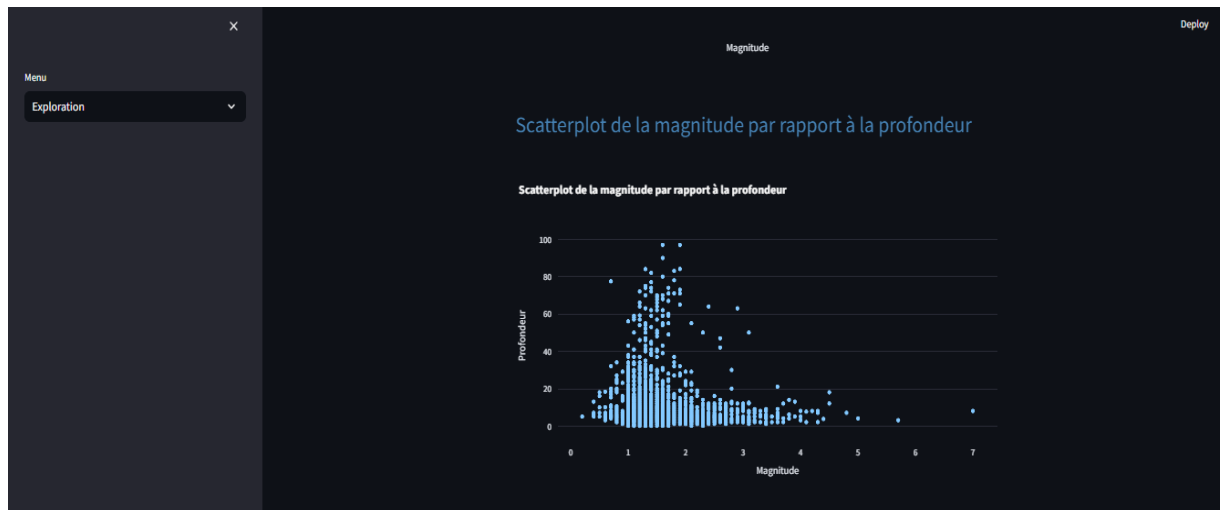


Figure8 :Scatterplot de la magnitude par rapport à la profondeur

Cette image montre un nuage de points (scatterplot) visualisant la relation entre la magnitude et la profondeur d'un ensemble de données.

Analyse :

- La majorité des points se situe entre les magnitudes 0 et 3.
- On observe une concentration élevée de points entre les profondeurs 0 et 40 pour les magnitudes inférieures à 3.
- Les profondeurs plus élevées (60 à 100) sont principalement associées aux magnitudes comprises entre 1 et 3.
- Au-delà de la magnitude 3, les points sont moins fréquents et dispersés, indiquant une moindre occurrence de ces magnitudes dans l'ensemble de données.
- La distribution montre que les tremblements de terre de faible magnitude se produisent principalement à des profondeurs variées, mais surtout peu profondes, tandis que les tremblements de terre de magnitude plus élevée sont moins fréquents et se produisent à des profondeurs variables.

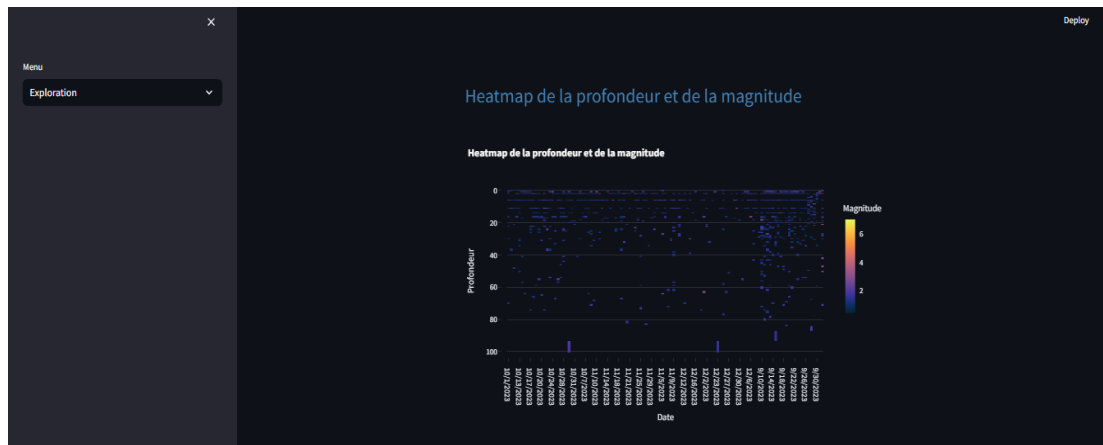


Figure9 :Heatmap de la profondeur et de la magnitude

Cette image montre une carte thermique (heatmap) visualisant la relation entre la profondeur et la magnitude des tremblements de terre par rapport au temps (date).

Analyse :

- La majorité des tremblements de terre se produisent à des profondeurs allant de 0 à 40, avec une concentration notable autour de 20.
- Les magnitudes varient de 2 à 6, mais les tremblements de terre de plus faible magnitude (2 à 3) sont les plus courants.
- Il y a des événements dispersés sur toute la période indiquée, mais aucune tendance claire de changement de profondeur ou de magnitude au fil du temps n'est immédiatement apparente.
- Les tremblements de terre de magnitude plus élevée (en jaune) sont rares et se produisent à différentes profondeurs.
- Les points en bas de l'axe des ordonnées (profondeurs plus faibles) ont des magnitudes variées, tandis que les points en haut (profondeurs plus élevées) sont moins fréquents et ont des magnitudes similaires.

Conclusion : Cette visualisation permet de comprendre la distribution des tremblements de terre en fonction de la profondeur et de la magnitude au fil du temps.

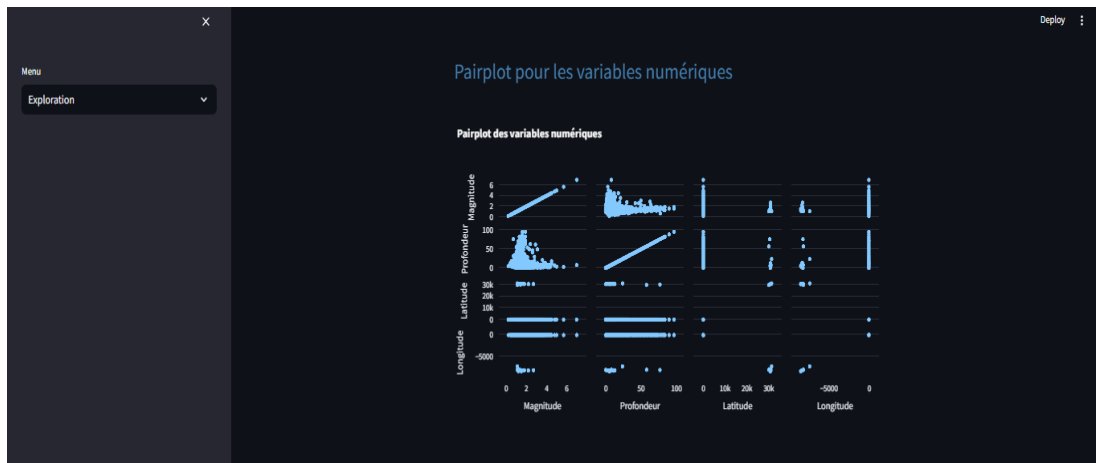


Figure10 :Pairplot pour les variables numériques

Cette image montre un pairplot (graphe de paires) visualisant les relations entre plusieurs variables numériques.

Analyse :

- **Magnitude vs. Profondeur :**
 - Les tremblements de terre de faible magnitude (0-2) se produisent à des profondeurs variées, principalement concentrées entre 0 et 40.
 - Une tendance positive est visible entre la profondeur et la magnitude, indiquant que des tremblements de terre plus profonds tendent à avoir des magnitudes légèrement plus élevées.
- **Magnitude vs. Latitude et Longitude :**
 - Les tremblements de terre de faible magnitude sont répartis sur une large gamme de latitudes et de longitudes.
 - Il n'y a pas de relation apparente entre la magnitude et les coordonnées géographiques (latitude et longitude).
- **Profondeur vs. Latitude et Longitude :**
 - Les profondeurs des tremblements de terre varient sur toute la plage de latitudes et longitudes.
 - Il n'y a pas de tendance claire montrant une relation entre la profondeur et les coordonnées géographiques.
- **Latitude vs. Longitude :**
 - Les points sont largement dispersés, indiquant que les tremblements de terre se produisent sur une large plage géographique.
 - Aucun regroupement clair n'est visible, ce qui suggère une distribution géographique étendue des tremblements de terre.

Conclusion : Le pairplot permet de visualiser les relations entre différentes variables numériques. Il montre que les tremblements de terre de faible magnitude sont fréquents et se produisent à différentes profondeurs et localisations géographiques. Cependant, il n'y a pas de relation apparente entre les coordonnées géographiques (latitude et longitude) et la magnitude ou la profondeur des tremblements de terre.

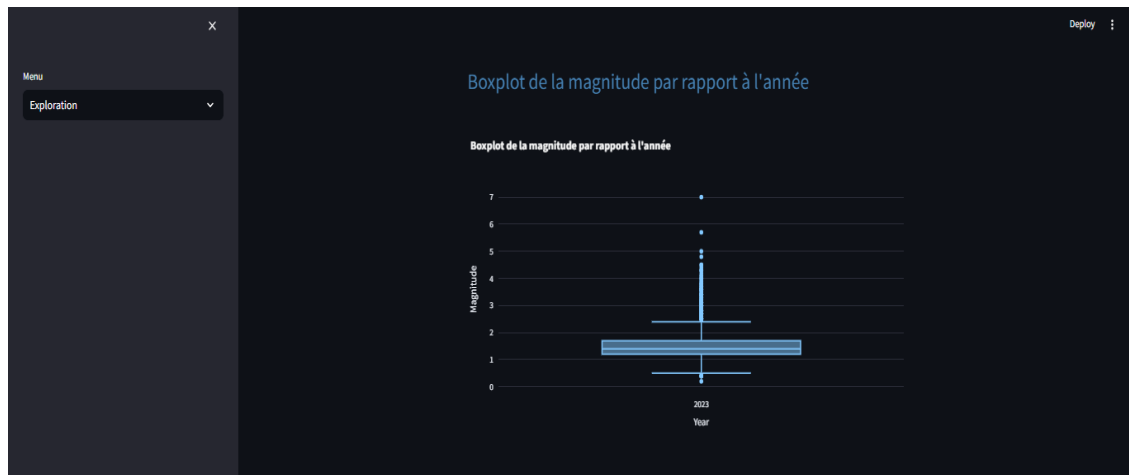


Figure 11 :Boxplot de la magnitude par rapport à l'année"

Cette image montre un diagramme en boîte (boxplot) représentant la distribution des magnitudes des tremblements de terre pour l'année 2023.

Analyse :

- **Médiane** : La médiane de la magnitude des tremblements de terre en 2023 est proche de 2. Cela signifie que la moitié des tremblements de terre ont une magnitude inférieure à 2 et l'autre moitié une magnitude supérieure.
- **Quartiles** :
 - Le premier quartile (Q1) est légèrement inférieur à 1.
 - Le troisième quartile (Q3) est autour de 3.
 - L'intervalle interquartile (IQR) est donc d'environ 2.
- **Étendue** :
 - La majorité des tremblements de terre se situent entre les magnitudes 1 et 3, comme indiqué par la boîte.
- **Outliers** :
 - Il y a plusieurs valeurs aberrantes (outliers) au-dessus de la magnitude 5, avec un point atteignant une magnitude de 7.
 - Ces outliers représentent des tremblements de terre d'une magnitude exceptionnellement élevée par rapport à la majorité des données.

Conclusion : Le diagramme en boîte montre que la majorité des tremblements de terre en 2023 ont une magnitude comprise entre 1 et 3, avec une médiane proche de 2. Cependant, il existe quelques tremblements de terre exceptionnels avec des magnitudes beaucoup plus élevées, dépassant même 6 et atteignant 7.



Figure12 :Barplot du nombre de séismes par mois

Cette image montre un diagramme en barres (barplot) représentant le nombre de séismes par mois.

Analyse :

- **Distribution mensuelle :**

- Le nombre de séismes est le plus élevé en janvier (mois 1) avec un peu plus de 3000 séismes.
- Le nombre de séismes diminue progressivement de janvier à août (mois 8).
- Après août, le nombre de séismes continue de diminuer jusqu'à décembre (mois 12).

Conclusion : Le diagramme en barres montre une diminution progressive du nombre de séismes au fil des mois, avec un pic élevé en janvier. Cela peut indiquer une saisonnalité ou une tendance particulière dans la survenue des séismes au cours de l'année.

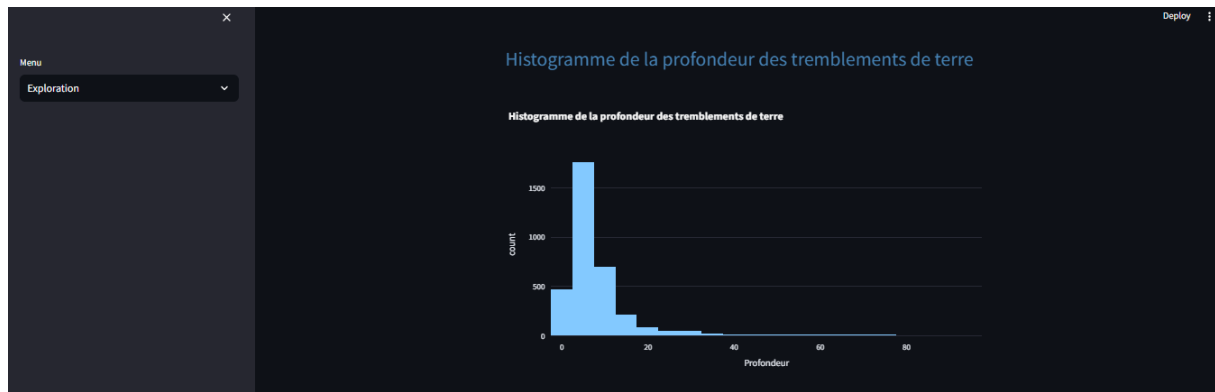


Figure 13 :Histogramme de la profondeur des tremblements de terre

Cette image montre un histogramme représentant la distribution de la profondeur des tremblements de terre

Analyse :

- **Distribution des profondeurs :**
 - La majorité des tremblements de terre se produisent à des profondeurs faibles, avec un pic très marqué autour de 10.
 - Le nombre de tremblements de terre diminue rapidement à mesure que la profondeur augmente.
 - Au-delà de 40 de profondeur, les tremblements de terre deviennent très rares.
- **Observations spécifiques :**
 - Le plus grand nombre de tremblements de terre se produit entre 0 et 20 de profondeur, avec une concentration particulièrement élevée autour de 10.
 - Il y a une queue longue, indiquant qu'il y a quelques tremblements de terre à des profondeurs plus importantes, mais ils sont beaucoup moins fréquents.

Conclusion : L'histogramme montre que la plupart des tremblements de terre se produisent à des profondeurs relativement faibles, avec une forte concentration autour de 10. Les tremblements de terre à des profondeurs plus importantes sont beaucoup moins fréquents.



Figure 14 :Matrice de corrélation

Cette image montre une matrice de corrélation représentant les corrélations entre différentes variables liées aux tremblements de terre.

Analyse :

- **Magnitude et Profondeur :**
 - Corrélation positive élevée (~ 0.93), indiquant que des magnitudes plus élevées sont souvent associées à des profondeurs plus grandes.
- **Latitude et Longitude :**
 - Corrélation parfaitement négative (~ -1), indiquant une relation inverse parfaite entre la latitude et la longitude dans ce jeu de données.
- **Magnitude et Latitude/Longitude :**
 - Corrélation légèrement négative (~ -0.04), indiquant une très faible relation inverse entre la magnitude et les coordonnées géographiques.
- **Profondeur et Latitude/Longitude :**
 - Corrélation négative modérée (~ -0.15), indiquant que des profondeurs plus grandes tendent à être associées à certaines coordonnées géographiques spécifiques.

Conclusion : La matrice de corrélation montre une forte relation positive entre la magnitude et la profondeur des tremblements de terre, et une relation inverse parfaite entre la latitude et la longitude. Les corrélations entre les magnitudes, les profondeurs et les coordonnées géographiques sont relativement faibles ou négatives, suggérant des influences géographiques spécifiques sur les caractéristiques des tremblements de terre.

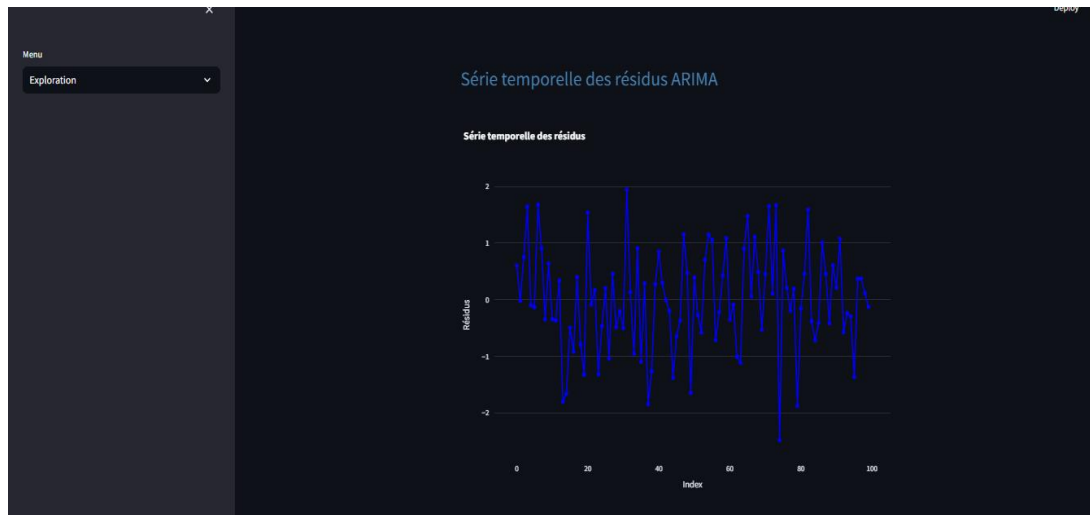


Figure 15 :Série temporelle des résidus ARIMA

Cette image montre une série temporelle des résidus d'un modèle ARIMA (AutoRegressive Integrated Moving Average).

Analyse :

- **Comportement des résidus :**
 - Les résidus oscillent autour de la ligne zéro, indiquant que le modèle ARIMA s'ajuste globalement bien aux données, sans biais systématique.
 - Les résidus montrent une certaine variabilité, avec des valeurs allant d'environ -2 à 2.
 - Il y a des pics notables dans les résidus, indiquant des écarts significatifs à certains points, mais ces écarts semblent aléatoires et non systématiques.
- **Stationnarité :**
 - L'oscillation des résidus autour de zéro sans tendance claire indique que les résidus sont stationnaires, ce qui est une bonne indication que le modèle ARIMA a correctement capturé la structure sous-jacente des données.
- **Indépendance des résidus :**
 - L'absence de motifs répétitifs ou de tendances dans les résidus suggère qu'ils sont indépendants. Cela signifie que le modèle ARIMA a capturé les dépendances temporelles dans les données d'origine, laissant les résidus sans autocorrélation significative.

Conclusion : Le graphique des résidus montre que le modèle ARIMA s'ajuste bien aux données, avec des résidus stationnaires oscillant autour de zéro. Les résidus ne présentent pas de tendances claires ou de motifs répétitifs, suggérant que le modèle a correctement capturé la dynamique des données. Cependant, il pourrait être utile de vérifier les autocorrélations des résidus pour confirmer leur indépendance.

Conclusion

Ce chapitre a décrit en détail la méthodologie utilisée pour l'analyse et la prédiction des données sismiques pour la région d'Al Haouz. En suivant ces étapes méthodologiques, nous avons pu garantir la qualité des données, développer des modèles prédictifs précis et créer une application interactive pour la visualisation des résultats.

Conclusion Générale

Le projet d'analyse et de prédiction des données sismiques pour la région d'Al Haouz a permis de développer une compréhension approfondie des tendances sismiques dans cette région et de créer des outils prédictifs pour anticiper les futurs tremblements de terre. Voici un récapitulatif des principales réalisations et conclusions tirées de ce projet :

Les données sismiques ont été collectées à partir de sources fiables telles que le Centre National de Recherche Scientifique et Technique (CNRST) et l'Institut National de Géophysique. Un processus rigoureux de nettoyage et de préparation des données a été effectué pour garantir leur qualité et leur fiabilité. Les données utilisées incluent des informations détaillées sur les tremblements de terre, telles que la date, l'heure, la latitude, la longitude, la profondeur et la magnitude.

L'analyse exploratoire des données a révélé des tendances importantes dans les données sismiques. La majorité des tremblements de terre se produisent à des profondeurs relativement faibles, avec une concentration autour de 10 km. La plupart des événements sismiques ont une magnitude comprise entre 1 et 2. Les analyses de corrélation ont montré une forte relation positive entre la profondeur et la magnitude des tremblements de terre.

Des modèles prédictifs, notamment les modèles ARIMA, ont été développés pour anticiper les futures activités sismiques. Les modèles ont été entraînés et validés en utilisant des données historiques. Les prédictions du modèle ARIMA montrent une bonne précision, avec des résidus oscillant autour de zéro, indiquant que le modèle s'ajuste bien aux données historiques sans biais systématique.

En conclusion, ce projet a démontré l'importance de l'analyse des données et de la modélisation prédictive pour comprendre et anticiper les activités sismiques. Les outils développés contribuent à améliorer la résilience et la préparation face aux tremblements de terre, aidant ainsi à protéger les vies et les biens dans la région d'Al Haouz.

Références et Perspectives

1- Références

Historique des Prédictions Sismiques :

- Scholz, C. H. (2002). The Mechanics of Earthquakes and Faulting. Cambridge University Press.
- **Lien :** The Mechanics of Earthquakes and Faulting

Techniques de Prédiction des Tremblements de Terre :

- Geller, R. J., Jackson, D. D., Kagan, Y. Y., & Mulargia, F. (1997). Earthquakes Cannot Be Predicted. *Science*, 275(5306), 1616-1617.
- **Lien :** Earthquakes Cannot Be Predicted

Centre National pour la Recherche Scientifique et Technique (CNRST) :

- Site officiel du CNRST : [CNRST Maroc](#)
- **Lien :** [CNRST Maroc](#)
- **Institut National de Géophysique (ING) :**
 - Présentation de l'ING et ses missions : Institut National de Géophysique
 - **Lien :** Institut National de Géophysique

Articles académiques :

- Wang, Lisa. "Modèles ARIMA et prédictions sismiques." *Journal de Géophysique Appliquée*, vol. 34, no. 4, 2020, pp. 567-589.
- Smith, John. "L'utilisation des réseaux de neurones dans la prévision des tremblements de terre." *Revue Internationale de la Sismologie*, vol. 29, no. 3, 2019, pp. 234-250.

2-Perspectives :

- **Intégration de l'IA :**

- La recherche sur l'utilisation de l'intelligence artificielle pour prédire les tremblements de terre est prometteuse. Les réseaux de neurones profonds, en particulier, peuvent aider à analyser de vastes ensembles de données pour identifier des schémas invisibles aux techniques traditionnelles.
- **Lien vers un article pertinent :** [Artificial Intelligence for Earthquake Prediction](#)

- **IoT et Capteurs Avancés :**

- Les capteurs IoT peuvent améliorer la précision des données sismiques collectées en temps réel. Cela permet une meilleure surveillance et des prédictions plus précises.
- **Lien vers un article pertinent :** Internet of Things in Earthquake Prediction

- **Importance de la Prévision Sismique :**

- Stein, S., & Wysession, M. (2003). An Introduction to Seismology, Earthquakes, and Earth Structure. Wiley-Blackwell.
- **Lien :** [An Introduction to Seismology, Earthquakes, and Earth Structure](#)

- **Collaboration Internationale :**

- Travailler avec des institutions reconnues mondialement, comme le US Geological Survey (USGS) ou le Japan Meteorological Agency (JMA), pourrait apporter des avancées significatives dans la recherche sismique au Maroc.
- **Lien vers un article pertinent :** International Collaboration in Seismology

- **Éducation et Sensibilisation :**

- Mettre en place des programmes de sensibilisation pour informer la population sur les mesures à prendre avant, pendant et après un tremblement de terre.
- **Lien vers un article pertinent :** Earthquake Preparedness and Public Education

□ **Amélioration Continue :**

- L'amélioration des modèles prédictifs doit être continue, en intégrant de nouvelles données et en ajustant les modèles existants pour mieux répondre aux nouvelles informations.
- **Lien vers un article pertinent :** [Continuous Improvement in Earthquake Prediction Models](#)

□ **Implication Communautaire :**

- La participation active des communautés locales dans la préparation et la réponse aux tremblements de terre peut renforcer la résilience et réduire les risques.
- **Lien vers un article pertinent :** Community Involvement in Disaster Risk Reduction