



RAPPORT DU STAGE OPÉRATEUR

Conception d'un outil de prévision des ventes des produits au sein des stations de service et les projets inhérents

Auteur :
Saad TAZROUTE

Encadrant :
Mr. Adil RADI

30 août 2019

Remerciements

J'adresse mes remerciements, dans un premier temps, à toute l'équipe pédagogique de l'**Ecole Centrale Casablanca** qui veille à la qualité de la formation au sein de l'école.

Aussi, je saisis cette occasion pour remercier profondément les responsables et le personnel du département Retail de la société Vivo energy pour leur bienveillance et leur convivialité.

Je remercie également Mr.**Reda BADAWI** pour son accueil chaleureux, et pour l'acceptation de ma candidature, m'ouvrant ainsi la porte pour cette belle expérience.

Je tiens à remercier vivement mon encadrant, Mr.**Adil RADI**, pour son engagement, ses conseils et le partage de son expertise. C'est grâce à sa confiance que j'ai pu réussir ma première expérience professionnelle.

Ensuite, je tiens à adresser mes remerciements les plus sincères à mon professeur, Mr.**Fouad RIANE**, pour ses cours très enrichissants et surtout celui de gestion d'entreprises qui m'était très utile, dans la mesure où je me suis basé sur ce dernier lors de la réalisation de ce projet. Sans oublier le cours de gestion des opérations -surtout la partie prévisions et planification auquel je me suis référé lors de ma période de stage comme source de documentation- enseigné par Mr.**Abdelghani BOURASS**.

Table des matières

1	Introduction	1
1.1	Contexte :	1
1.2	Buts et utilités	2
1.3	Schématisation de la mission :	3
1.3.1	Intitulé la mission :	3
1.3.2	Problématique :	3
1.3.3	Objectif :	3
1.3.4	Hypothèses :	3
1.3.5	Démarche :	3
1.4	Planning du rapport :	4
2	Présentation du périmètre du projet :	5
2.1	Contexte général :	5
2.2	Présentation du secteur pétrolier au Maroc :	6
2.2.1	Historique du secteur pétrolier :	6
2.2.2	Aperçu du secteur pétrolier au Maroc :	6
2.3	Présentation de VIVO ENERGY MAROC	8
2.3.1	Présentation, Mission, Vision, Buts, Objectifs et DAS	8
2.3.2	Analyse externe de Vivo Energy Maroc :	11
2.3.3	Analyse SWOT de VEM :	14
3	La théorie des méthodes des prévisions :	17
3.1	Les sources de données :	17
3.2	La typologie de la demande :	18

3.3	Les méthodes de prévision :	19
3.3.1	Les méthodes qualitatives :	19
3.3.2	Les méthodes quantitatives :	20
3.3.3	Le lissage exponentiel :	21
3.3.4	Méthode ARIMA et SARIMA :	23
3.4	Mesure de performance :	23
3.5	Simulation de prévision :	24
3.6	Evaluation de l'erreur de chaque méthode appliquée sur la base de données mensuelles :	27
3.7	Prévision des ventes grâce à la méthode ARIMA appliquée sur la base de données journalières :	31
4	Analyse et bilan du stage du stage :	33
	Conclusion	35
5	Annexe 1 : Description des bases de données utilisées	39
6	Annexe 2 : Méthodes classiques	41
7	Annexe 3 : Méthode ARIMA	51

Table des figures

1.1	La société CDC Arkhinéo termine l'année 2018 en ayant dépassé deux milliards de documents archivés sur sa plateforme. La société franchit également le cap du Peta-octet (1 million de Go) de documents confiés	2
2.1	Consommation par région en million tonne	7
2.2	L'évolution de lu logo Shell à travers le temps	9
2.3	Quelques indicateurs économiques du Maroc	11
2.4	Vue globale sur le marché marocain des produits pétroliers	13
2.5	Bientôt la fin des énergies fossiles?	14
2.6	SWOT analysis of VEM	15
3.1	Le processus à suivre pour avoir des prévisions	19
3.2	Description de la base de données mensuelles	26
3.3	Description de la base de données journalières	26
3.4	Base de données mensuelles	26
3.5	Base de données journalières	26
3.6	Evaluation de l'erreur de la méthode de la moyenne	27
3.7	Evaluation de l'erreur de la méthode des moyennes mobiles Résultat très satisfaisant, méthode à prendre en considération	28
3.8	Evaluation de l'erreur du lissage simple	28
3.9	Evaluation de l'erreur du lissage double(méthode de Holt)	29
3.10	Evaluation de l'erreur du lissage triple version (Holt-Winter)	30
3.11	Evaluation de l'erreur de la méthode ARIMA	30
3.12	Les erreurs liés aux méthodes présentées	31
3.13	Prévision à l'aide de la méthode ARIMA	31

5.1	Base de données mensuelles	39
5.2	Base de données journalières 01/01/2019 .. 26/07/2019	39

Liste des sigles et acronymes

VEM	<i>Vivo Energy Morocco</i>
GPL	<i>Le gaz de pétrole liquéfié</i>
Forecasting	Prévision
ARIMA	<i>Auto Regressive Integrated Moving Average</i>
SES	<i>Simple exponential smoothing</i>
RMSE	<i>Root mean square error</i>
MAE	<i>Mean Absolute Error</i>
ME	<i>Mean Error</i>

Chapitre 1

Introduction

1.1 Contexte :

La big data fait partie des enjeux majeurs pour l'avenir. Les entreprises restant à la marge prennent le risque de se voir dépasser par la concurrence. Le défi, et notamment pour les PME-PMI, est que toutes aient un système d'information avec une capitalisation sur la donnée client. Dans ce cadre, l'objectif est par exemple de prévoir un meilleur ciblage client ou encore des prévisions de vente affinées.

Dans le cadre de solutions issues du Big Data, la relation client peut connaître des transformations très importantes. Grâce à l'analyse des données massives issues des réseaux sociaux et du web, il est désormais possible d'approfondir de manière substantielle notre connaissance des clients : ce qu'ils aiment ou pas dans notre produit, leur façon de l'utiliser, les caractéristiques de consommation par zone géographique, etc. Il est possible aussi de prédire à quel point un projet inhérent peut-il être rentable en réalisant une prévision de ventes à l'aide d'un outil de prédiction.

La transformation numérique au sein des entreprises constitue un défi sociétal majeur pour la majorité des entreprises. Pour VEM, c'est Optima qui accompagne cette migration vers le numérique.

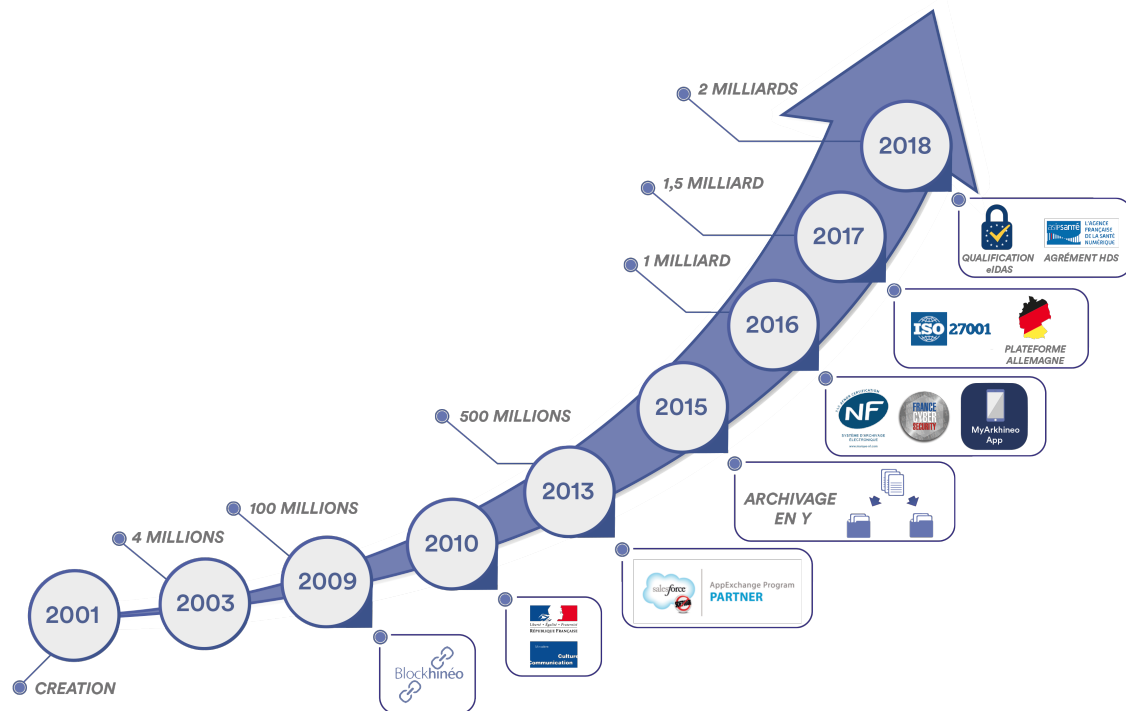


FIGURE 1.1 – La société CDC Arkhinéo termine l’année 2018 en ayant dépassé deux milliards de documents archivés sur sa plateforme. La société franchit également le cap du Peta-octet (1 million de Go) de documents confiés

1.2 Buts et utilités

Il y a plusieurs raisons pour l’utilisation d’un système de recommandation :

- **Augmentation des revenus** : Cette fonctionnalité est probablement la plus importante dans un contexte industriel (Commercialisation des nouveaux produits au sein des shops dans les nouvelles stations de service). L’objectif ici est de vendre plus de produits qu’on aurait pu vendre . Pour atteindre cet objectif, le système prédit les articles qui seront les plus rentables pour l’entreprise.
- **Etude de rentabilité des projets inhérents** : Le but de cette fonction est d’inciter les chefs de projets de construction de stations de service à faire une étude au préalable de la possibilité de vente d’un produit, remettre en question la différence de ventes entre les produits, mieux comprendre pourquoi un tel ou tel produit ne se vend, et ainsi pouvoir travailler sur ces axes d’amélioration.
- **Prendre en compte les préférences du client** : Si le système fonctionne correctement et qu’il est conçu dans un environnement approprié, il peut augmenter la conscience de l’entreprise vis à vis des préférences et des choix du client et alors augmenter sa satisfaction.

1.3 Schématisation de la mission :

1.3.1 Intitulé la mission :

Exploitation des données fournies par l'entreprise pour élaborer une prévision des ventes des différents produits dans les stations de service.

1.3.2 Problématique :

Le département retail essaye d'augmenter son efficacité dans le processus de lancement de nouveaux projets (nouvelles stations de service). Pour pouvoir exceller dans l'opération du lancement, il faut avoir une idée concernant les ventes des produits sur le long terme.

1.3.3 Objectif :

Conception d'un outil de prévision des ventes des produits au sein des projets de stations de service et les projets inhérents.

1.3.4 Hypothèses :

1. Les données fournies par l'entreprise sont des données qui connaissent des anomalies.
2. On suppose que toutes les stations sont du même niveau d'importance, et ont le même rythme de ventes.
3. On négligera les fluctuations des données, qui sont principalement liés à des crises (par exemple, le boycott des produits de l'entreprise Afrika Gaz en Juin 2017), ou à l'emplacement géographique de la station.
4. On supposera qu'il n'y aura pas d'instabilités (guerre, grève, boycott ..) sur le court terme (2019/2021).

1.3.5 Démarche :

1. Avoir une vision globale sur le marché pétrolier mondial et surtout le marché marocain pour estimer le poids de l'entreprise.
2. Réalisation d'un diagnostic stratégique de l'entreprise.
3. Se documenter sur les séries chronologiques.
4. Comprendre les approches majeures utilisées des méthodes de prévisions.
5. Se familiariser avec les bibliothèques utilisées dans les méthodes de prévision basées sur les séries chronologiques.

6. Implémentation de différentes méthodes.
7. Proposer la méthode la plus adaptée au problème de ventes dans les stations de service.

1.4 Planning du rapport :

Dans le Chapitre II, on va présenter le périmètre du projet, et on passera en revue sur la situation de l'entreprise (organisation, vision, mission, valeurs).

Dans le chapitre III, on va présenter les concepts de base de prévision basée sur les séries chronologiques ainsi que leurs modélisations mathématiques, et on abordera principalement la méthode du lissage exponentiel. On implémentera finalement tout ce qu'on a présenté sur python et on discutera les résultats obtenus. On finira par une conclusion globale du travail.

Chapitre 2

Présentation du périmètre du projet :

Cette partie est consacrée à passer en revue la situation de l'entreprise (organisation, vision, mission, valeurs). Je vais d'abord aborder le contexte général du sujet, ensuite je vais présenter la structure des secteurs pétroliers mondiaux et marocains, et celle de VEM. Par la suite, j'effectuerai un diagnostic stratégique de l'entreprise et sur la base des données issues de cette analyse, je déduirai les facteurs clés de succès, la capacité stratégique de l'entreprise et une synthèse SWOT.

2.1 Contexte général :

En Décembre 2015 l'État Marocain a levé la main sur la fixation du prix du gasoil et de l'essence. Comme résultat trivial, le secteur pétrolier est devenu de plus en plus attrayant au Maroc. Depuis cette date, les entreprises ont beaucoup investi afin de garantir le meilleur service à la clientèle. Cette révolution dans le secteur pétrolier m'a beaucoup intéressé, dans la mesure où la privatisation joue un rôle moteur dans l'évolution de n'importe quel secteur.

Dans le cadre de ma formation à l'Ecole Centrale Casablanca, j'ai pris la décision de passer mon stage opérateur au sein d'une entreprise à renommée mondiale dans le domaine d'énergie. Je me suis fixé en fin de compte sur l'entreprise Vivo Energy (Franchise de Shell en Afrique) et en particulier dans le département "Retail" de cette entreprise.

La mission du stage m'était une vraie opportunité pour me familiariser avec les termes techniques liés à tout ce qui est gestion de données, prévision, et avoir le premier contact avec tout ce qui est technique.

Cette mission m'a attiré particulièrement, puisque j'ai souhaité découvrir si ce type de métier m'intéressait vu que je m'oriente dans ma formation et dans ma carrière professionnelle vers le domaine de la science des données.

L'entreprise Vivo Energy, créée très récemment en 2011, s'est fait différencier avec succès par ses services innovants. Pour cette raison, entre autres, j'ai voulu intégrer ses équipes pour pouvoir découvrir l'atmosphère dans une entreprise qui était considérée comme meilleur recruteur au Maroc en 2018.

Dans le présent rapport, je m'emploierai à élaborer une étude de cas de Vivo Energy Maroc. Ainsi, je préciserai la structure organisationnelle, l'organisation du service dans lequel j'ai effectué mon stage. Ensuite, je présenterai ma mission de stage, la démarche que j'ai suivie pour atteindre les objectifs listés par mon encadrant, les résultats et les difficultés que j'ai rencontrés lors de ma mission.

2.2 Présentation du secteur pétrolier au Maroc :

2.2.1 Historique du secteur pétrolier :

La demande mondiale de pétrole, et principalement d'essence, a fortement augmenté dans la région hors OCDE (Organisation de coopération et de développement économiques) [Figure 1] principalement en Chine, en Inde, au Pakistan et au Vietnam.

Malgré la suppression des subventions, la mise en place de nouvelles taxes et les politiques en cours pour accroître les biocarburants dans plusieurs pays de la région asiatique. La croissance de la demande d'essence s'est poursuivie. En outre, la croissance mondiale de la demande de distillats moyens s'est également accélérée au premier trimestre, sous l'impulsion des récentes hausses en Chine et en Inde.

Dans le bassin atlantique, de nouveaux règlements sont entrés en vigueur pour établir des zones de contrôle du milieu marin, ce qui a entraîné le remplacement du mazout résiduel par du diesel marin dans la demande d'équilibrage, stimulant ainsi la consommation de carburant diesel. [Figure 1]

2.2.2 Aperçu du secteur pétrolier au Maroc :

Le Maroc est le pays d'Afrique du Nord où les premières recherches de pétrole, déjà anciennes, firent naître l'espoir d'y découvrir d'importantes réserves. Cet espoir, un moment renforcé à une date récente par le succès des prospections entreprises sur certains points de la zone des plaines atlantiques, a été déçu : Le Maroc ne paraît pas devoir devenir un grand producteur de pétrole. Avec une extraction, en 1963, de 146 000 tonnes de pétrole, le Royaume chérifien est un très modeste producteur qui ne peut satisfaire à sa consommation annuelle sans cesse croissante.

Être un pays déficitaire en matière d'énergie, était la cause majeure qui a poussé l'état marocain à instaurer un système de subvention, or ceci grève et déséquilibre la balance

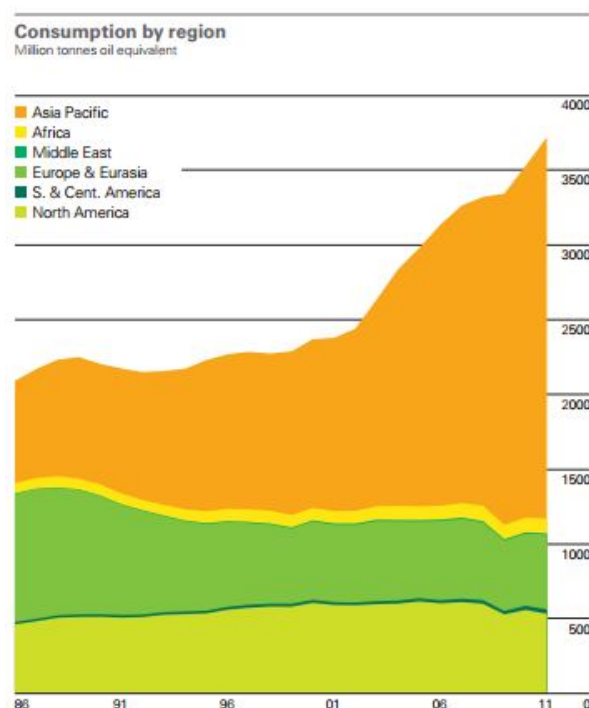


FIGURE 2.1 – Consommation par région en million tonne

des paiements, d'où la nécessité de libéraliser le secteur des produits pétroliers. Afin de régulariser l'approvisionnement du marché des produits de base et protéger le pouvoir d'achat des citoyens, l'état marocain a instauré dans le début des années quarante un système de subvention qui pourra maîtriser le niveau des prix et le flux d'importation et d'exportation.

La compensation a joué pendant plusieurs années un rôle primordial dans le soutien et le développement de plusieurs secteurs notamment des secteurs agroalimentaires tel que les secteurs laitier, huilier et sucrier, ou encore industriels tel que les produits pétroliers et le ciment.

La caisse de compensation est dotée par le budget de l'Etat des fonds nécessaires pour mener cette mission qui est devenue de plus en plus délicate par égard aux sommes faramineuses des subventions à déboursier pour ces produits et qui ont atteint des niveaux alarmants au cours de ces dernières années. Les produits pétroliers monopolisaient la plus grande part des dépenses de subvention avec plus de 86% du montant total, la charge était passée de près d'un milliard de dirhams en 2003 pour atteindre 41 milliards de dirhams en 2011 et 48 milliards en 2012.

Cette charge de compensation a alourdi de plus en plus les finances publiques au dépend d'autres investissements qui pourraient stimuler l'économie et développer le pays. La part

de l'enveloppe dédiée à la compensation dans le PIB est passée de 1% en 2003 à 6,5% en 2012. Cette situation impactait négativement l'équilibre des finances publiques en creusant le déficit budgétaire qui est passé de 2,2% en 2009 à 7,3 % en 2012 et qui a dépassé les 7,5%, au terme du 1er semestre 2013.

La figure 2 ci-dessous permet de quantifier l'évolution de la charge de compensation qui connaît des pics lors des crises alimentaires mondiales en 2008 et à partir de 2010 pour exploser en 2011, année tumultueuse des printemps arabes, avec le doublement des subventions ayant pour enjeu le maintien de la paix sociale dans un contexte politique et économique incertains.

2.3 Présentation de VIVO ENERGY MAROC

Ayant choisi d'user au maximum les connaissances acquises au sein de l'école Centrale de Casablanca, j'ai choisi de mettre en exergue dans cette partie, un exemple d'analyse stratégique pour permettre de tirer les conclusions adéquates par la suite en se basant principalement sur le cours de gestion d'entreprise enseigné par Mr. Fouad RIANE. C'est ainsi qu'en allant de la présentation de VEM à son analyse interne, je tâcherai au mieux de décrire le mode de fonctionnement de Vivo Energy Maroc.

2.3.1 Présentation, Mission, Vision, Buts, Objectifs et DAS

Présentation et mission de VIVO ENERGY MAROC

Identité visuelle de Shell

L'identité visuelle de Shell a connu des changements au fil des années afin de devenir le logo actuel qui fut l'œuvre du designer industriel Raymond Loewy.

Codétenue par Vitol, Helios Investment Partners et Shell, le Groupe possède un réseau de plus de 1800 stations-service dans 15 pays et exporte des lubrifiants vers plusieurs autres pays africains.

Filiale du groupe, la mission de Vivo Energy Maroc (VEM) est donc de « **contribuer à l'avenir de l'Afrique** ». C'est dans ce sens que tout est mis en œuvre afin de faciliter la vie des clients et de rendre leur expérience chez VEM plus pratique, plus agréable et plus bénéfique.



FIGURE 2.2 – L'évolution de lu logo Shell à travers le temps

Vision de Vivo Energy Maroc

Vivo Energy a pour ambition de « *Devenir la société d'énergie la plus respectée d'Afrique* ». Pour la société, cela signifie faire affaire dans les règles et donner la priorité à la santé, la sûreté, la sécurité et l'environnement.

VEM souhaite faire affaire dans les règles, parce que c'est le meilleur moyen de réussir et d'assurer la croissance de l'entreprise. Cela n'est donc pas simplement fait pour des raisons d'éthique, d'altruisme ou par manque d'ambition.

Les normes appliquées et respectées en Afrique sont les mêmes normes strictes qui s'appliquent sur les marchés en aval les plus sévèrement régulés au monde. Au sein de l'entreprise, il existe une tolérance zéro en matière de corruption, pratiques dangereuses, négligence environnementale ou concurrence déloyale.

Buts de Vivo Energy Maroc

Ils se déclinent surtout à travers la composante de la vision de VEM qui souhaite donner la priorité à la santé, la sécurité, la sûreté et l'environnement (HSSE). Pour VEM, cela signifie :

- Objectif « Zéro accident » : donner la priorité à la sécurité. Cela est très remarquable aussi bien au sein de l'entreprise qu'à l'extérieur .
- Investir dans l'avenir : pour la société, les communautés et les pays dans lesquels VEM exerce ses activités.
- Apporter une valeur ajoutée aux collaborateurs : autonomiser, reconnaître, former et récompenser.

— Prendre soin de l’environnement, des communautés et du personnel.

Les deux précédents points lui ont d’ailleurs valu d’être élu « Meilleur employeur 2018 » en mars dernier.

Domaines d’activités stratégiques

Le groupe SHELL est impliqué dans tous les aspects de la filière énergétique : prospection de gisements, extraction de pétrole et de gaz naturel, raffinage, pétrochimie, vente (B2B et détail). Pour ce qui est de Vivo Energy Group mais surtout de Vivo Energy Maroc, c’est surtout le dernier domaine d’activité stratégique de SHELL qui résume son activité principale : la distribution et la commercialisation des produits de la marque SHELL. Pour mener à bien ses ambitions, Vivo Energy Maroc a subdivisé son activité principale comme suit :

- **Offre sur le réseau**
- **Offre commerciale**

Les principaux domaines d’activités stratégiques de la filiale de Vivo Energy au Maroc sont au nombre de six et permettent de couvrir en totalité la réalisation de son activité principale. Ce sont :

1. **Le transport** : qui inclut l’importation et la logistique importante qui permet de ravitailler les différents centres du Maroc ;
2. **Les produits financiers** : afin de contrôler les aspects financiers qu’impliquent les activités de l’entreprise.
3. **Le carburant** : il constitue l’un des principaux DAS puisque qu’il permet de satisfaire une clientèle très diversifiée tout en assurant un chiffre d’affaires important. Des motos aux navires en passant par les véhicules à quatre roues, les avions et les équipements miniers, le carburant de marque Shell prend différentes formes à fin de pouvoir créer de la valeur pour le client qui utilise n’importe quel type de véhicule.
4. **Le lubrifiant** : On distingue les huiles moteur Shell HELIX pour automobiles, les huiles moteur Shell RIMULA pour camions et poids lourds et les huiles moteur Shell ADVANCE pour motos.
5. **Le GPL** : Vivo Energy Maroc est leader des distributeurs de GPL dans le pays, sous la marque Butagaz. L’entreprise distribue du butane sous forme de bouteille pour les segments domestique, petit commercial et industriel, mais aussi en vrac aux clients des secteurs industriel et commercial (secteurs de la céramique, de la mise en boîte, du verre, de l’industrie chimique et des hôtels, etc.).
6. **Boutiques et services** Au-delà des shops Shell Select, les stations-services généralisent progressivement la fourniture d’une gamme complète d’installations et de services dont le lavage de voitures, des baies de graissage, un service de restauration rapide.

2.3.2 Analyse externe de Vivo Energy Maroc :

1. *Analyse du macro-environnement :*

L'analyse externe passe comme à l'accoutumée par **une analyse PESTEL** : Politique, Economique, Socio-démographique, Technologique, Ecologique et Légal.

(a) **Environnement politique :**

Le contexte politique qui est actuellement vécu au Maroc s'inscrit dans le cadre d'une monarchie constitutionnelle assurant un certain équilibre au niveau du royaume. En effet, le Maroc jouit d'une grande stabilité politique et ne rencontre pas de tensions internes majeures. Il est l'un des pays les plus stables de l'Afrique du Nord, ce qui lui a d'ailleurs permis de développer de bonnes relations au niveau international ainsi que sur le continent Africain.

(b) **Environnement économique :**

Le Maroc a renoué avec la croissance économique depuis le début des années 2000, qui était alors corrélée au secteur agricole et à la conjoncture mondiale. Son économie penche beaucoup plus de nos jours vers le secteur tertiaire et la production agricole ne parvient toujours pas à satisfaire les besoins des populations. Par ailleurs, et sous pression du FMI avec lequel le Maroc est lié depuis août 2012, le gouvernement a réintroduit une indexation partielle du prix de certains produits pétroliers et supprimé la subvention à l'essence et au fuel industriel en justifiant cette décision par les contraintes liées à la concurrence et par la chute des cours du pétrole à l'international, visant ainsi la libéralisation du marché et l'instauration de la concurrence entre les distributeurs en faveur des consommateurs. Ce qui a nettement amélioré le climat des affaires et contribué à l'attractivité de l'économie marocaine.

Indicateur (2016)	Valeur
PIB	103.606 Milliards USD
Taux de croissance	1,5 %
PIB par habitant	2832 \$USD
Taux de chômage	9,4%

FIGURE 2.3 – Quelques indicateurs économiques du Maroc

(c) **Environnement socio-démographique :**

Le Maroc a connu une croissance démographique très rapide à travers le temps, pour compter officiellement 34.996.167 habitants en 2017 avec plus de 84000 ressortissants étrangers sur son territoire. Son emplacement géographique fait du Royaume un creuset de cultures diversifiées, par les arrivées successives. Les

communautés y vivent en harmonie, chacune respectant les différences des autres, vu la diversité de sa population berbère, arabe, musulmane, juive, africaine, et andalouse. Cette mosaïque d'ethnies lui confère une grande richesse culturelle.

(d) **Environnement technologique :**

Le secteur des nouvelles technologies au Maroc connaît plusieurs mutations en raison de la dynamique du marché, de l'apparition de nouvelles innovations et de l'adoption de nouvelles politiques à l'image du plan Maroc numérique. A vrai dire, L'accès aux technologies de bases telles que l'ordinateur ou internet par la population rurale surtout reste encore difficile, même s'il y a une amélioration notable. Ainsi, en contribuant à la croissance économique et à la construction d'une société de la connaissance, les technologies de l'information ont pu devenir un vecteur de création de richesses à part entière dans le pays, faisant dorénavant partie de stratégies nationales voire internationales.

(e) **Environnement écologique :**

L'équilibre écologique du Royaume est entaché par de nombreux facteurs, climatiques et industriels surtout. A cet effet, le gouvernement s'est lancé dans plusieurs réformes et a mené des réflexions autour de plusieurs projets dans ce sens ; l'interdiction des sachets plastiques en est un exemple palpable. Ces mesures sont aujourd'hui appuyées par la charte nationale pour l'environnement, qui a pour ambition de porter le développement durable au rang de valeur fondamentale et vise à susciter une prise de conscience écologique au niveau national.

(f) **Environnement légal :** Les principaux textes qui régissent le secteur pétrolier au Maroc sont explicitées dans le Dahir portant loi n°1-72-255 du 18 moharrem 1393 (22 février 1973), tenant compte des modifications apportées par la loi n°4-95 le complétant.

2. *Les 5+1 forces de Porter*

(a) **Le pouvoir de négociation des clients,**

Vivo Energy Maroc dispose d'une clientèle très diversifiée et répandue sur tout le territoire. En effet, de Laayoune à Nador, des particuliers aux grandes entreprises, chacun y trouve son compte. Elle fidélise sa clientèle grâce à la qualité de son carburant, la diversité de ses offres et ses nombreuses promotions qui permettent d'attirer de temps à autre un achalandage important.

(b) **Le pouvoir de négociation des fournisseurs :**

VEM dispose d'une multitude de fournisseurs, tous départements confondus. De la construction au carburant en passant par l'électricité, ce sont des secteurs diversifiés qui apportent leur soutien à la bonne marche de l'activité de VEM. Notons que le seul département Retail mobilise une vingtaine de fournisseurs tandis que le département Ingénierie en a 51 à son actif.

(c) **Intensité concurrentielle :**

Le marché de la distribution des produits pétrolier est un marché fortement concurrentiel. Vivo Energy Maroc occupe la deuxième place dans le réseau de distribution des produits pétroliers. En 2016, elle est précédée par Afriquia et suivie par Total Maroc. Ces deux entreprises constituent donc les concurrents les plus proches de VEM.

(d) **La menace d'entrants potentiels sur le marché :**

Dans le secteur pétrolier, on peut observer de nombreuses barrières à l'entrée, telles que la réalisation d'économies d'échelle grâce à l'atteinte d'une taille critique, les forts besoins en capitaux, la notoriété, les avantages du premier entrant, ... Ces barrières freinent fortement l'arrivée de nouveaux entrants sur le marché du pétrole. Il est donc difficile pour une entreprise venant d'être créée d'envisager concurrencer des géants tel que Vivo Energy, Total et surtout Afriquia. En parallèle à cela, le fait que le pétrole ne soit pas durable rend le secteur peu attractif. Malgré tous ces aspects, le fait que cette activité soit très lucrative peut à l'inverse attirer de nouvelles entreprises sur le marché, surtout depuis que l'État a levé la main sur la fixation du prix du gasoil et de l'essence en décembre 2015, les opérateurs ne cessent de s'enrichir.

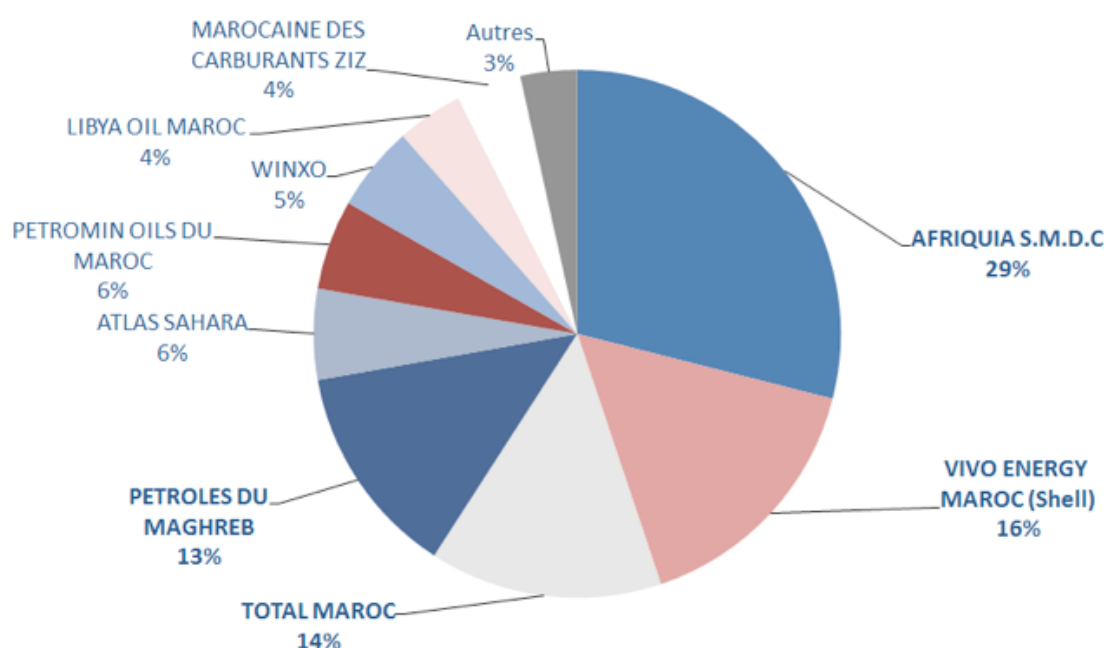


FIGURE 2.4 – Vue globale sur le marché marocain des produits pétroliers

(e) **Menace des produits de substitutions :**

Actuellement, on ne peut pas véritablement parler de produits de substitution. En revanche, face à la pénurie probable de pétrole qui pourrait se manifester d'ici quelques décennies, des énergies alternatives sont en voie de développement. Les alternatives qui pourront éventuellement substituer les produits pétroliers sont principalement les énergies renouvelables. Si la menace actuelle est faible, il est probable que d'ici quelques années elle devienne de plus en plus forte.

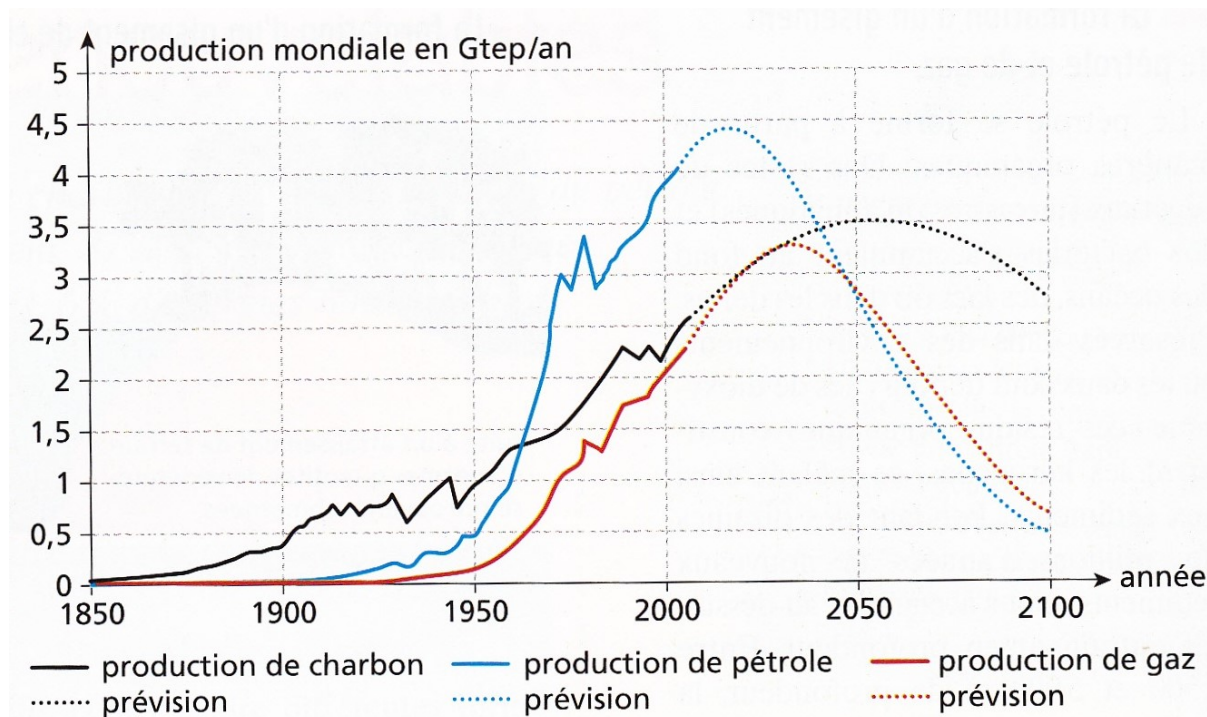


FIGURE 2.5 – Bientôt la fin des énergies fossiles ?

2.3.3 Analyse SWOT de VEM :

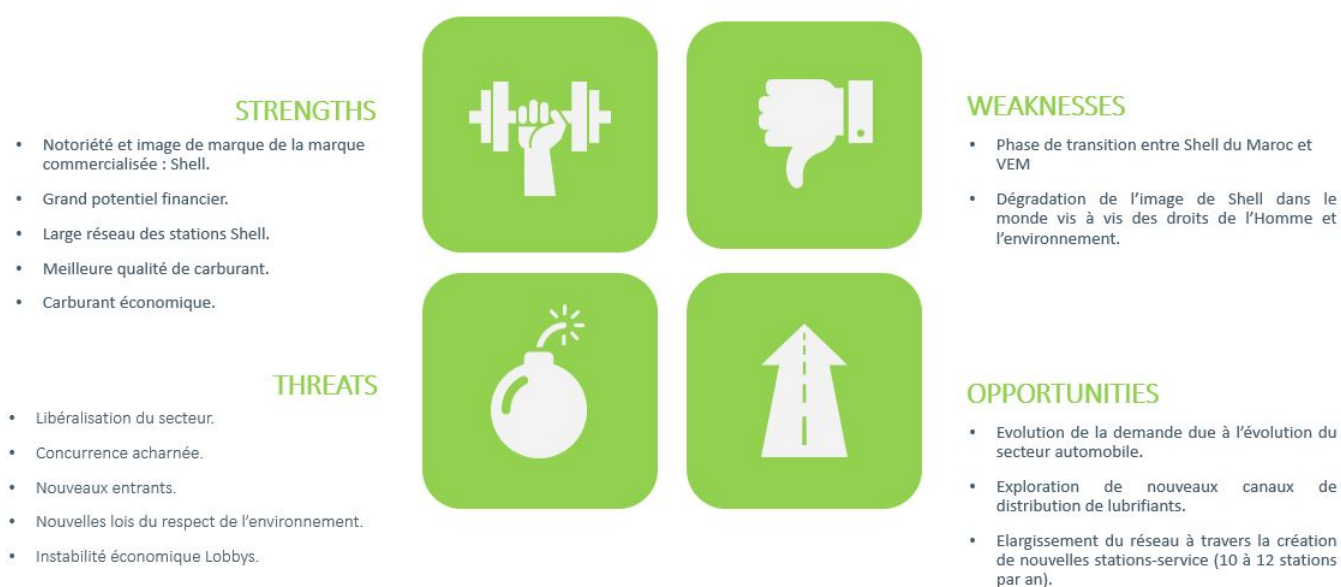


FIGURE 2.6 – SWOT analysis of VEM

Chapitre 3

La théorie des méthodes des prévisions :

La prévision de la demande est définie comme une projection dans le futur de la demande, connaissant un certain nombre de conditions environnementales. L'activité de prévision est le point de départ de la planification.

Précisons que, dans un environnement instable - comme c'est le cas aujourd'hui- la prévision est difficile. Toutefois, mieux vaut prévoir même avec incertitude que de ne pas le faire !

Les prévisions n'étant par nature que rarement exactes, il est important d'obtenir la meilleure qualité de prévision possible. Pour cela, différents modèles de prévisions ont été élaborés, en particulier dans le domaine des statistiques. Les questions qui se posent alors sont : Quels sont les modèles existants, quels sont leurs avantages et inconvénients respectifs ? Comment choisir le modèle de prévision le plus performant pour une base de données fournie par l'entreprise ? C'est à ces réflexions que nous essayons d'apporter des solutions dans ce chapitre.

3.1 Les sources de données :

Les sources de données correspondent aux deux familles de méthodes de prévision : la première fondée sur des données relevées dans le passé que l'on modélise pour faire une projection dans le futur, et l'autre, purement prédictive, établie par des experts interrogés.

La source privilégiée de données est un historique de données concernant un produit. Cette base permet d'effectuer une prévision si, évidemment, on estime qu'il existe un lien entre l'évolution de la demande passée (données enregistrées) et celle de la demande à prévoir.

Les autres sources de données sont constituées par les études de marché, les avis d'experts, le suivi des commerciaux, les enquêtes auprès des clients... Mais ces données sont plus délicates à manipuler et interpréter. En revanche, elles constituent un complément sûr à un historique, sans compter que si l'on ne dispose pas d'historique, c'est la seule source utilisable.

Ceci va être détaillé davantage dans ce qui suit.

3.2 La typologie de la demande :

L'étude des phénomènes économiques a distingué depuis longtemps divers types d'évolution, qui peuvent éventuellement se combiner :

- * **La tendance** (T_t) : correspond une variation lente s'effectuant dans un sens déterminé qui se maintient pendant de longues périodes.

- * **La composante saisonnière** (S_t) : correspond à des variations s'effectuant régulièrement au cours de la semaine, du mois, du trimestre, etc. Elle tient aux saisons, des habitudes sociologiques et aux rythmes de l'activité humaine. (comme par exemple les mois du Ramadan pour les grandes surfaces).

- * **Les fluctuations accidentelles** : sont des mouvements erratiques, de fréquences élevées, présentant une allure générale plus ou moins stable. Elles résultent des influences, que toutes sortes d'événements exercent sur la grandeur en cause, si nombreuses que l'on a renoncé à les étudier toutes dans le détail.

Certaines séries statistiques présentent l'un ou l'autre de ces mouvements à l'état pur. Mais la plupart d'entre elles ont une allure plus complexe.

On peut donc distinguer 5 grands types de demande :

Demande constante si elle oscille statistiquement autour d'une valeur moyenne constante dans le temps, la moyenne de $D = f(t)$ est une droite horizontale.

Demande à tendance s'il y a oscillation autour d'une valeur croissante ou décroissante dans le temps, $D = f(t)$ est une droite de pente positive ou négative.

Demande saisonnière si elle présente des variations nettement plus importantes, en hausse et en baisse, d'une manière périodique. Il peut s'agir d'un pic de la demande en hiver (lié à la neige par exemple) ou en été (vacances) mais il peut aussi s'agir de variations saisonnières plus subtiles.

Demande saisonnière et à tendance si les pics et les creux sont disposés autour d'une droite non horizontale.

Demande erratique si les valeurs sont totalement aléatoires dans le temps.

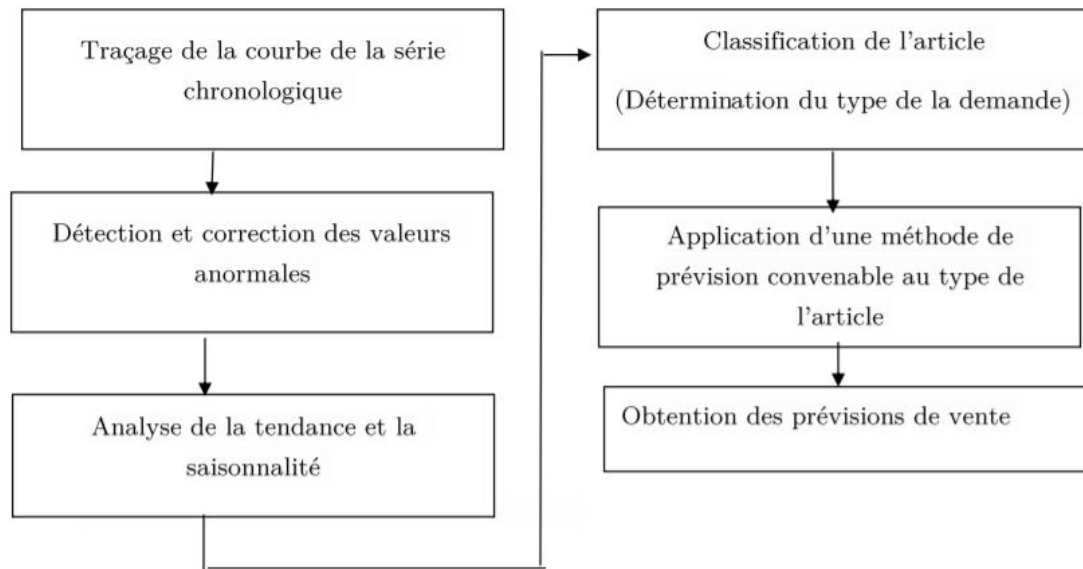


FIGURE 3.1 – Le processus à suivre pour avoir des prévisions

3.3 Les méthodes de prévision :

On distingue deux grands types de méthodes de prévision : les méthodes qualitatives et les méthodes quantitatives. Les techniques qualitatives font appel à une méthodologie non mathématique (mais elles peuvent impliquer des valeurs numériques). Les techniques quantitatives, au contraire, seraient fondées sur des modèles mathématiques. De plus, ces techniques sont dites intrinsèques si les données manipulées sont celles du produit considéré. Elles sont extrinsèques s'il s'agit de données appartenant des événements relatifs l'article mais qui ne le concernent pas directement.

3.3.1 Les méthodes qualitatives :

Les méthodes qualitatives utilisent des données subjectives. Le résultat dépend du jugement, de l'expérience et de l'expertise de ceux qui formulent les prévisions. Elles se basent principalement sur :

- * L'opinion des vendeurs.
- * L'opinion des consommateurs (enquête).
- * L'opinion d'experts.
- * L'opinion des cadres.

Les méthodes qualitatives les plus connues sont :

- * Les enquêtes auprès des consommateurs
- * Les panels d'experts
- * La méthode de Delphi
- * Les analogies historiques

Les méthodes qualitatives permettent, l'inverse des méthodes quantitatives, de prendre en compte des facteurs intangibles. De plus, lorsqu'il n'existe que très peu de données, comme par exemple lors de l'introduction d'un nouveau produit ou lors de la pénétration d'un nouveau marché.

Ces méthodes peuvent s'avérer particulièrement pertinentes. Cependant, le processus de consultation d'experts ou de consommateurs peut être long et coûteux. Il existe également un risque d'obtenir une prévision biaisée ou arbitraire. Enfin, la précision de ce type de prévision est limitée. La méthode de Delphi : consiste à interroger des experts sur une question, et ce indépendamment les uns des autres afin de nier toute influence forte directe. Le coordinateur remet l'ensemble des réponses aux experts qui peuvent modifier et compléter leur proposition. Après deux ou trois cycles de ce type, on parvient à une proposition de consensus efficace, ou éventuellement à des divergences argumentées. Cette technique n'est pas adaptée à une prévision à court terme d'un article, mais au contraire à une décision de stratégie long terme.

3.3.2 Les méthodes quantitatives :

Les méthodes quantitatives sont des méthodes purement mathématiques et statistiques. Deux grandes familles de méthodes quantitatives existent : les familles chronologiques qui permettent de prévoir la demande en fonction de données historiques, et les méthodes causales appelées aussi méthodes associatives qui permettent d'établir des relations de cause effet entre certaines variables de l'environnement et la demande. Puisque dans notre cas, je n'avais accès qu'à l'historique des ventes, je me suis focalisé sur les méthodes temporelles seulement.

Les méthodes des séries chronologiques :

* **Méthode naïve** : La prévision pour la période suivante (période $t+1$) sera égale à la demande réelle de cette période. En d'autres mots, on utilise la valeur réelle de la dernière période comme une prévision.

* **Méthode de la moyenne** : Utilise simplement une moyenne de toutes les données passées comme une prévision.

$$P_t = \frac{1}{n} \sum_{i=1}^n d_k$$

* **La moyenne mobile** : La prévision pour la période suivante (période $t+1$) sera égale à la moyenne d'un certain nombre d'observations les plus récentes, chaque observation ayant la même importance (poids).

$$P_t = \frac{1}{n} \sum_{i=1}^n d_{t-k}$$

3.3.3 Le lissage exponentiel :

En général, le principe se base sur une moyenne pondérée dont le poids diminue de façon exponentielle au fur et à mesure que les données vieillissent. L'avantage c'est qu'on peut tenir compte des tendances et des saisonnalités. On va entamer davantage les méthodes suivantes vu que tout notre travail se base sur cette théorie.

Les méthodes de lissage exponentiel, basées sur les travaux de Holt et Winters, ne font plus l'objet de recherches. Si elles peuvent apparaître "faibles" théoriquement, elles sont faciles à mettre en œuvre, d'où leur utilisation populaire. Leur objectif est de prévoir à la date T une série temporelle à un horizon l à partir de T observations (X_1, \dots, X_T) .

Le lissage exponentiel simple (LES) :

Un lissage exponentiel simple est utilisé lorsque les données n'ont pas de tendance ni de saisonnalité. Le lissage exponentiel simple (LES) consiste à supposer que X_t est approximable autour de T par une constante a_T . Soit $\hat{X}_t(\ell)$ la prévision de $X_{T+\ell}$ à l'instant T .

La prévision par la méthode du LES est la suivante :

$$\forall \ell \in \mathbb{N}^* : \hat{X}_T(\ell) = \hat{X}_T(1) = \hat{X}_{T+1} = (1 - \alpha) \sum_{t=0}^{T-1} \alpha^t X_{T-t}$$

où $\alpha \in]0, 1[$ désigne le coefficient de lissage. On peut remarquer que :

Les réalisations ont d'autant plus de poids qu'elles sont récentes (en effet le poids α^t devant X_{T-t} décroît avec t).

Il n'est pas obligatoire de recalculer entièrement la prévision à chaque instant. On a en effet :

$$\begin{aligned}\hat{X}_{T+1} &= (1 - \alpha) X_T + \alpha \hat{X}_T \\ \text{D'où : } \hat{X}_{T+1} &= \hat{X}_T + (1 - \alpha) (X_T - \hat{X}_T)\end{aligned}$$

On peut ainsi calculer la lissée \hat{X}_t à n'importe quel instant t en fonction de celle à l'instant $t-1$.

Choix du coefficient de lissage On utilise un algorithme d'optimisation afin de minimiser la quantité suivante :

$$\sum_{t=1}^{T-1} \left(X_{t+1} - \hat{X}_{t+1} \right)^2 = \sum_{t=1}^{T-1} \left(X_{t+1} - (1 - \alpha) \sum_{i=0}^{t-1} \alpha^i X_{t-i} \right)^2$$

Cette quantité est la somme des écarts quadratiques entre les réalisations et les prévisions de la série temporelle.

La méthode de lissage exponentiel ne nécessite que deux données pour l'appliquer (la demande réelle la plus récente et les prévisions les plus récentes).

Une caractéristique attrayante de cette méthode est que les prévisions faites avec Ce modèle inclura une partie de chaque élément de la demande historique. De plus, ces valeurs historiques de la demande seront pondérées différemment, les données plus anciennes recevant des pondérations plus faibles.

N.B : Plus α est proche de 1 plus les observations récentes influent sur la prévision, à l'inverse un α proche de 0 conduit à une prévision très stable prenant en compte un passé lointain.

Le lissage exponentiel double (méthode de Holt) :

La méthode de Winter est utilisée lorsque les données ont un caractère saisonnier, mais aucune tendance.. La formule pour l'estimation de la valeur lisse est la suivante : Le lissage exponentiel double (LED) consiste à supposer que X_t est approximable au voisinage de T par une droite : $a_T + (t - T) b_T$.

Où a_T et b_T sont des solutions d'un problème d'optimisation qu'on n'aura pas besoin de savoir pour le moment.

Les formules de mise à jour sont :

$$\begin{cases} \hat{a}_T = \hat{a}_{T-1} + \hat{b}_{T-1} + (1 - \alpha^2) (X_T - \hat{X}_{T-1}(1)) \\ \hat{b}_T = \hat{b}_{T-1} + (1 - \alpha)^2 (X_T - \hat{X}_{T-1}(1)) \end{cases}$$

Le paramètre de la méthode du LED est $\alpha \in]0, 1[$.

La prévision par la méthode du LED est la suivante :

$$\forall \ell \in \mathbb{N}^* : \hat{X}_T(\ell) = \hat{a}_T + \ell \hat{b}_T .$$

Le lissage exponentiel triple (méthode de Holt-Winter) :

Dans ce cas, les données doivent être à la fois tendanciellles et saisonnières.

La méthode de Holt-Winters, consiste à supposer que X_t est approximable au voisinage de T par $a_T + (t - T)b_T + S_T$. En désignant par s la période du cycle saisonnier de la série temporelle, les formules de mise à jour sont :

$$\begin{cases} \hat{a}_T = (1 - \alpha) (X_T - \hat{S}_{T-s}) + \alpha (\hat{a}_{T-1} + \hat{b}_{T-1}) \\ \hat{b}_T = (1 - \beta) (\hat{a}_T - \hat{a}_{T-1}) + \beta \hat{b}_{T-1} \\ \hat{S}_T = (1 - \gamma) (X_T - \hat{a}_T) + \gamma \hat{S}_{T-s} \end{cases}$$

Les paramètres de la méthode de Holt-Winters sont α, β et γ (tous dans $]0,1[$). La prévision par la méthode de Holt-Winters est la suivante :

$$\begin{cases} \hat{X}_T(\ell) = \hat{a}_T + \ell \hat{b}_T + \hat{S}_{T+\ell-s} & \text{si } \ell \in \{1, \dots, s\} \\ \hat{X}_T(\ell) = \hat{a}_T + \ell \hat{b}_T + \hat{S}_{T+\ell-2s} & \text{si } \ell \in \{s+1, \dots, 2s\} \\ \dots \end{cases} .$$

3.3.4 Méthode ARIMA et SARIMA :

Dans les statistiques et économétrie , et en particulier dans l'analyse des séries chronologiques , une autorégressif moyenne mobile intégrée (ARIMA) modèle est une généralisation d'une moyenne mobile autorégressif modèle (ARMA). Ces deux modèles sont adaptés aux données de séries chronologiques , soit pour mieux les comprendre, soit pour prévoir les points futurs de la série (prévision). Les modèles ARIMA sont appliqués dans certains cas où les données indiquent une non-stationnarité , où une étape de différenciation initiale (correspondant à la partie "intégrée" du modèle) peut être appliquée une ou plusieurs fois pour éliminer la non-stationnarité.

3.4 Mesure de performance :

Les entreprises prouvent souvent de grandes difficultés à mesurer « objectivement » la qualité d'une prévision.

Dans la partie précédente, j'ai passé en revue les principaux modèles et méthodes de prévision. Malgré ce nombre considérable de modèles, aucune approche ne peut être jugée comme étant a priori la meilleure technique. C'est pourquoi nous allons voir maintenant les méthodes de validation et de suivis des modèles de prévisions.

Une fois un modèle défini, il va falloir le valider et mettre en place un suivi permettant d'obtenir des valeurs cohérentes et de vérifier ses qualités prévisionnelles pour éventuellement lui apporter des modifications afin de le rendre plus pertinent.

Il est donc important de choisir les bons. Il existe de très nombreux indicateurs permettant d'évaluer la qualité de la prévision, on présente ici les plus importants :

- **Erreur moyenne ou mean error (ME) :**

C'est la moyenne des erreurs de prévision sur les n périodes :

$$ME = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i}$$

- **Carré moyen d'écarts ou medium square error (MSE) :**

le carré moyen des écarts est préféré à l'écart absolu moyen (ME) car il permet de privilégier un système qui introduit plusieurs erreurs ayant de faibles écarts à un système introduisant des erreurs moins fréquentes mais avec des écarts plus importants.

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2$$

- **Ecart moyen absolu ou mean absolute deviation(MAE) :** $MAE = \frac{1}{n} \sum_{i=1}^n |e_i|$

- **Root mean square error (RMSE) :** $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2}$

Avec $e = Y_{observed} - Y_{predicted}$

3.5 Simulation de prévision :

Quel est le volume vendu du nouveau carburant dans les 12 premiers mois ? Quelle est la tendance de la demande pour les produits émergents.

Le lissage exponentiel nous promet la possibilité de jeter un coup d'œil sur l'avenir en construisant des modèles.

Une méthode plus avancée utilisée dans ce sujet est la méthode ARIMA qu'on va implémenter et comparer les résultats.

On va illustrer dans la suite comment construire Lissage exponentiel simple, Holt, et Holt-Winters modèles utilisant **Python** et **Statsmodels**. Ces méthodes sont déjà implémentées et sont prêtes à être utilisées.

On présentera le script utilisé pour les méthodes suivantes en **Annexe 2** :

- Méthode de la moyenne
- La moyenne mobile
- Lissage exponentiel simple
- Lissage exponentiel double
- Lissage exponentiel triple

On va présenter les résultats de chaque méthode par la suite.

Pour les données disponibles, grâce à l'encadrant j'ai pu avoir une base de données

à **fréquence mensuelle** de vente de 2 produits à partir du mois **Octobre de l'année 2016** jusqu'au mois de **Juillet 2019**. Une autre base de données plus prometteuse(en termes de nombre de données), est une base de données à **fréquence journalière** à partir de **01/01/2019** jusqu'au **26/07/2019**, qui contient les ventes dans plusieurs stations au Maroc des carburants. On s'intéressera à mener l'étude sur la station **AL JAZIRA** et plus précisément les ventes du gasoil durant toute cette période. Les deux bases de données utilisées sont présentées en **Annexe 1**. Par la suite on va consacrer une section afin de présenter les résultats d'une méthode plus avancée qui est la méthode **ARIMA**. Le script est présenté en **Annexe 3** .

Nombre de données	34
Moyenne	255,294118
Ecart type	74.611125
Minimum	99
Maximum	413

FIGURE 3.2 – Description de la base de données mensuelles

Nombre de données	207
Moyenne	407887.127971
Ecart type	67816.696478
Minimum	253937.6
Maximum	585661.3

FIGURE 3.3 – Description de la base de données journalières

Décomposition des séries temporelles :

On peut décomposer une série temporelle pour avoir une idée sur son comportement saisonnier et tendanciel.

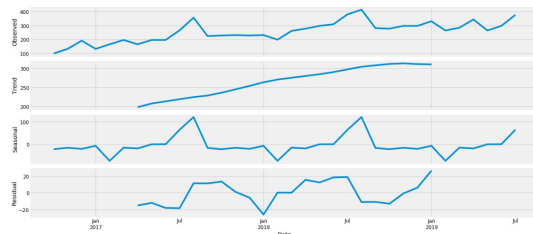


FIGURE 3.4 – Base de données mensuelles

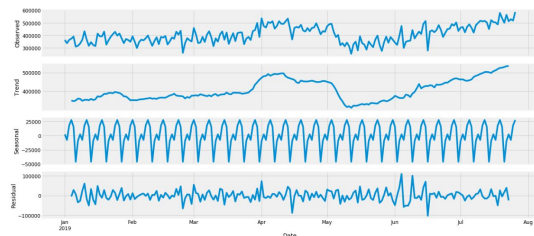


FIGURE 3.5 – Base de données journalières

3.6 Evaluation de l'erreur de chaque méthode appliquée sur la base de données mensuelles :

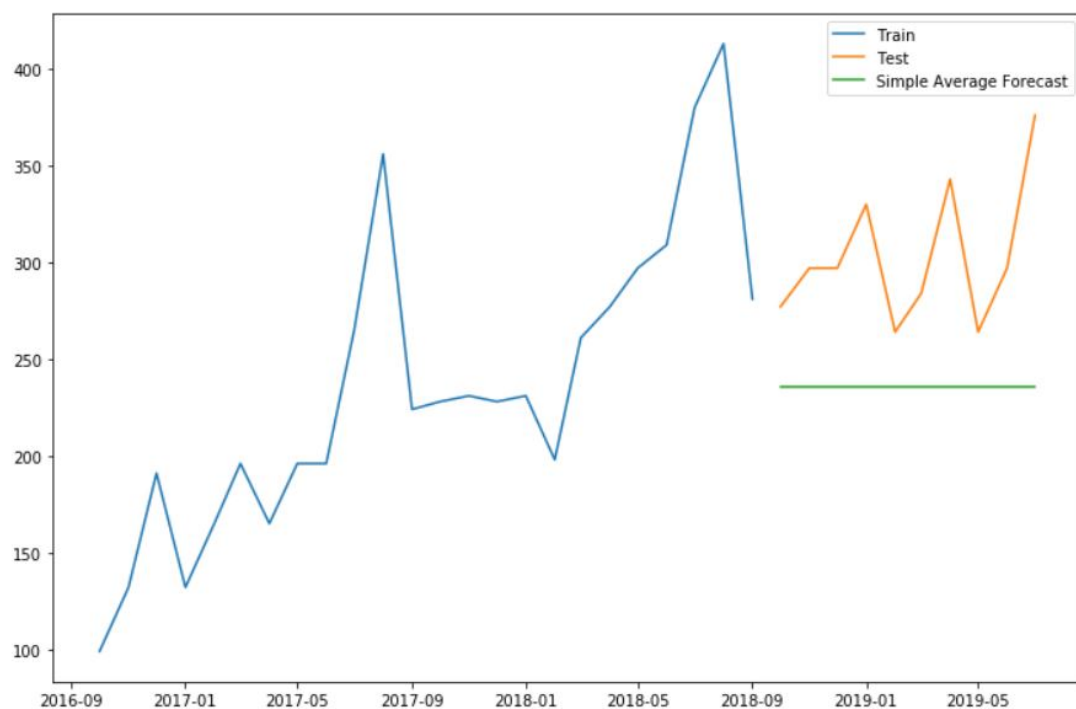


FIGURE 3.6 – Evaluation de l'erreur de la méthode de la moyenne
Le résultat obtenu est très loin de la réalité. Méthode à rejeter.

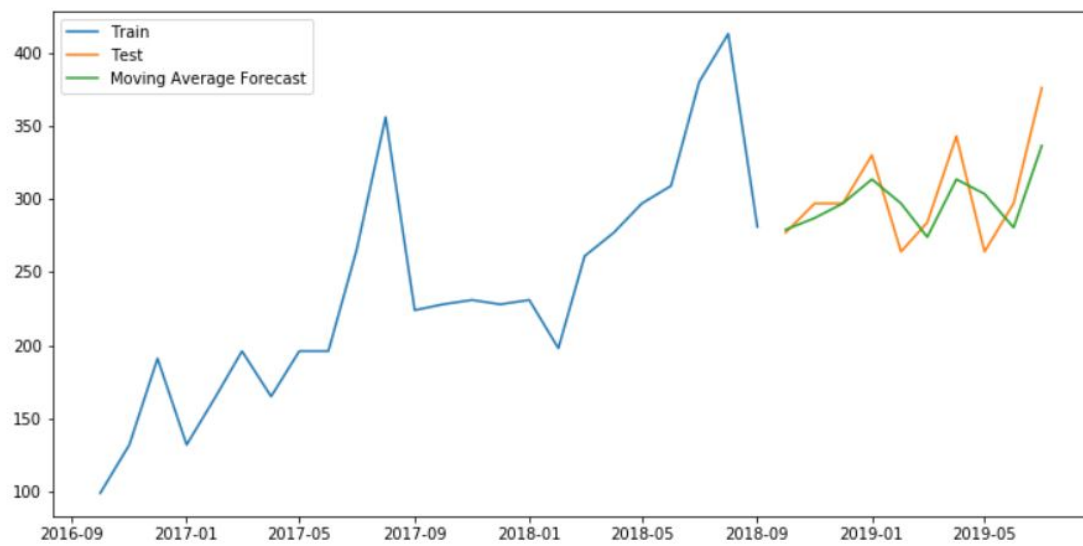


FIGURE 3.7 – Evaluation de l'erreur de la méthode des moyennes mobiles
Résultat très satisfaisant, méthode à prendre en considération



FIGURE 3.8 – Evaluation de l'erreur du lissage simple

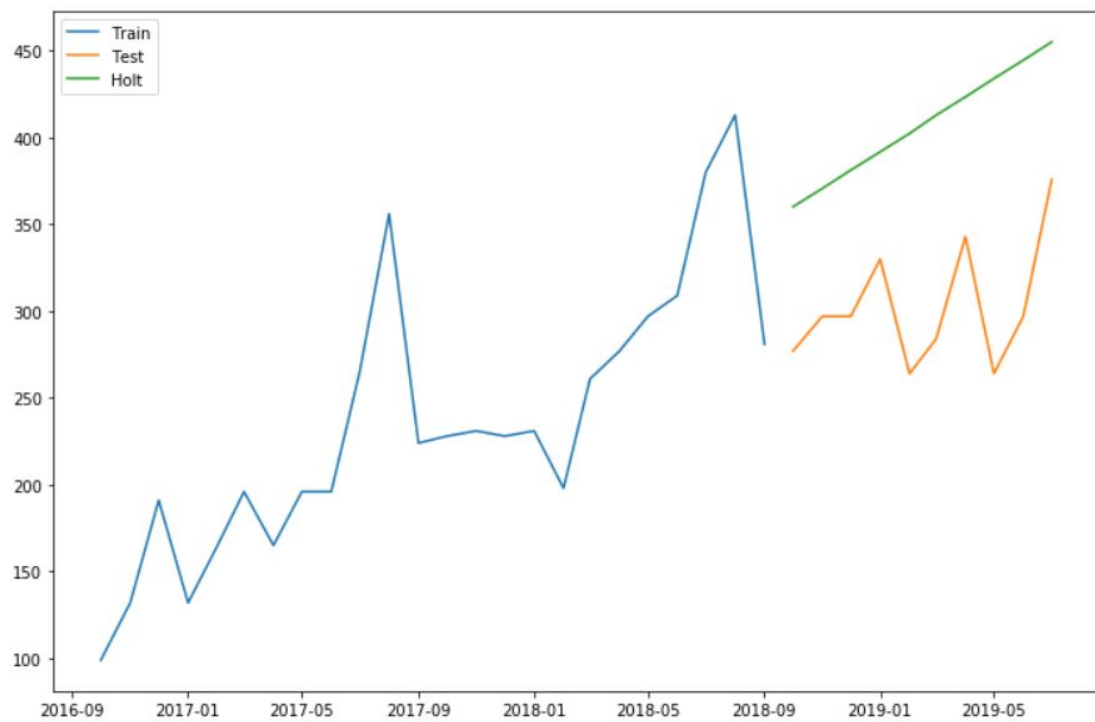


FIGURE 3.9 – Evaluation de l'erreur du lissage double(méthode de Holt)

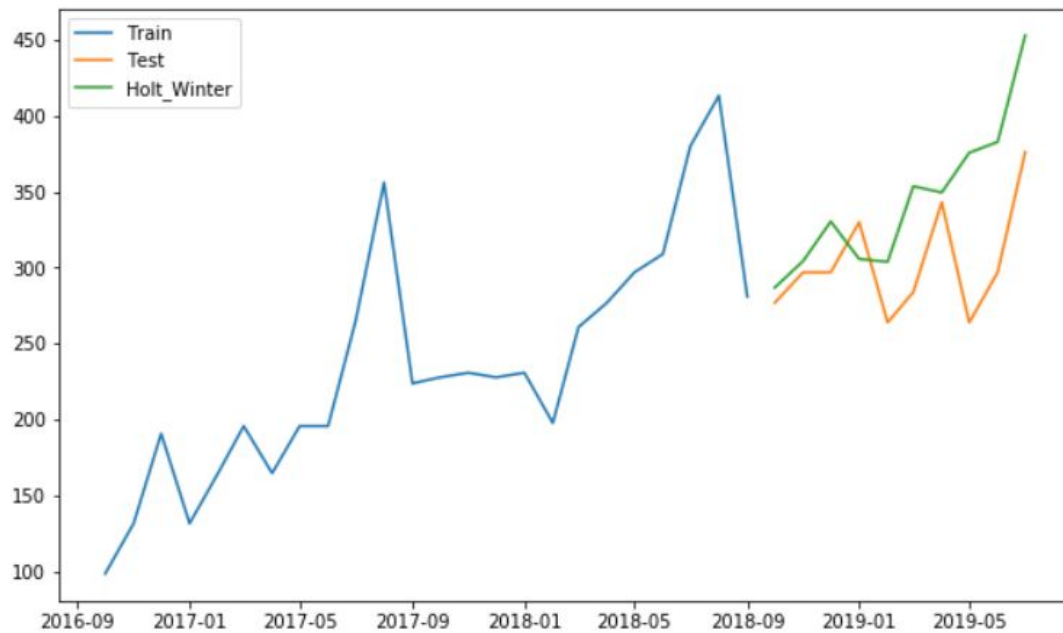


FIGURE 3.10 – Evaluation de l'erreur du lissage triple version (Holt-Winter)
Mauvais résultats puisqu'il s'éloigne de la réalité. Ceci peut être dû à la structure de la base de données fournie.

En effet chaque base de données a ses caractéristiques, et le choix du model reste lié principalement à la nature des données fournies. La méthode des moyennes mobiles est la seule méthode qui donne des résultats très satisfaisantes.

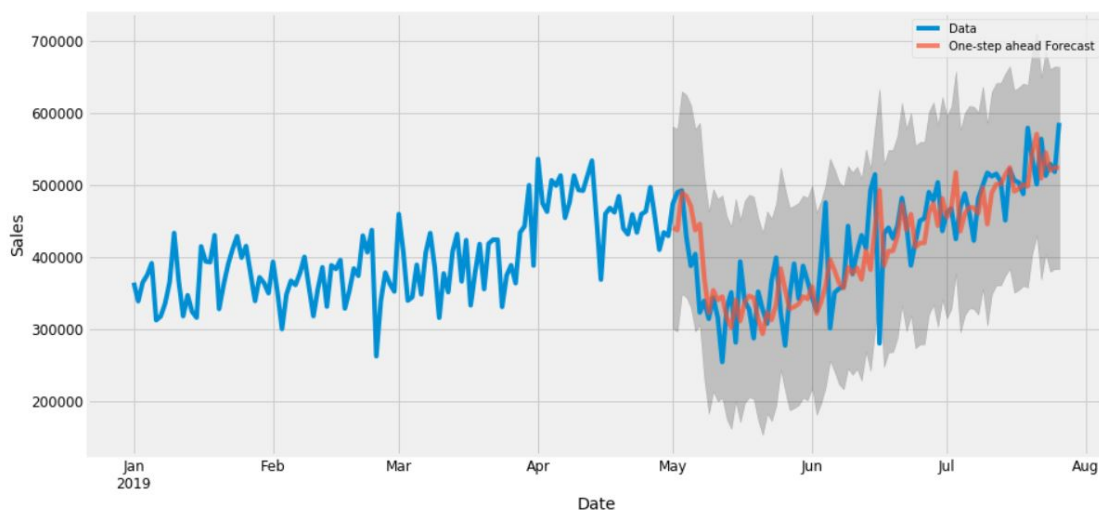


FIGURE 3.11 – Evaluation de l'erreur de la méthode ARIMA

Méthode utilisée	RMSE	MAPE
Simple average forecast	95.716	0.2133
Moving average forecast	44.142	0.0647
SES forecast	55.600	0.0825
Holt	130.666	0.3584
Holt Winter forecast	78.300	0.1577
ARIMA forecast	35,46	0.052

FIGURE 3.12 – Les erreurs liés aux méthodes présentées

On a bien constaté que la méthode la plus performante est une méthode non classique, l'utilisation de la méthode ARIMA donne des résultats plus satisfaisants.

On va appliquer cette méthode sur la base de données à fréquence journalière.

3.7 Prévision des ventes grâce à la méthode ARIMA appliquée sur la base de données journalières :

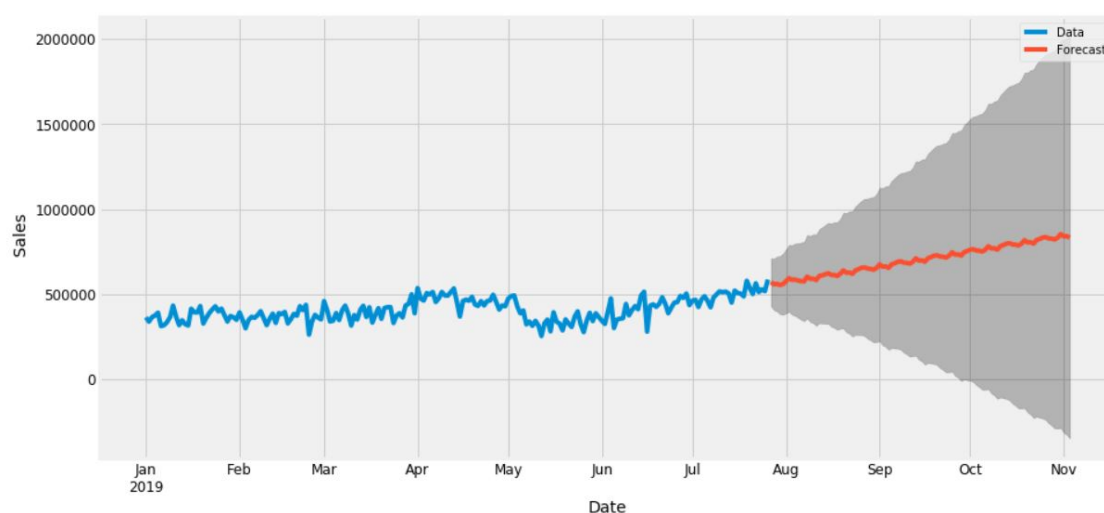


FIGURE 3.13 – Prévision à l'aide de la méthode ARIMA

Un fichier Excel contenant le résultat de cette prévision jusqu'au début du mois Novembre 2020 est fourni à l'encadrant comme résultat final.

Chapitre 4

Analyse et bilan du stage du stage :

Mon premier stage en entreprise n'était pas simplement une porte ouverte pour découvrir le monde de l'entreprise ou bien une mission à accomplir dans une période de 6 semaines, elle fut également une expérience riche d'acquis professionnels et personnels, surtout sur la partie communication avec le tuteur avec mes collègues.

Parmi les points majeurs que j'ai pu mettre en œuvre durant ces 6 semaines, on trouvera l'apprentissage rapide et efficace, car parmi tout les articles que je dois lire et synthétiser je dois les trier afin de pouvoir avancer sur ma mission tout en gardant le contact avec mon encadrant que je remercie vivement à cette occasion. Durant ce stage j'ai pu évoluer sur le volet technique en entamant de nouvelles bibliothèques soit sur Python et pouvoir les manipuler et ne pas rester bloquer sur une erreur. La capacité de trouver l'erreur et pouvoir la corriger en se référant soit aux forums ou au site fournisseur représente pour moi une capacité qu'un data scientist doit acquérir.

J'ai aussi suivi une formation en ligne afin de me documenter à propos des séries chronologiques, et j'ai réussi à avoir un certificat d'apprentissage.

Toutefois, passer une période dans une entreprise, nécessite suivre un code de vie qui organisera la vie en communauté au sein de la société. Personnellement, le projet enjeu, les réunions avec les tuteurs et les séances d'ADPL étaient très utiles pour moi puisque je n'ai rencontré aucun problème sur ce volet, cependant la majorité des stagiaires que j'ai rencontrés durant mes 6 semaines, ont rencontré des problèmes concernant la ponctualité et le code vestimentaire qu'il doivent respecter.

Ce stage m'a permis en premier lieu d'avoir une idée sur l'état d'avancement de ma formation au sein de l'école, dont je suis très satisfait, et aussi une idée sur la qualité de la formation au sein de l'école.

Conclusion et perspectives

Parmi les problèmes cruciaux auxquels sont confrontées les entreprises modernes, on trouve l'organisation du travail et des ressources productives et savoir à quel point l'approche suivie par l'entreprise est efficace, et pour remédier à ceci on peut faire appel à un système de prévision afin de prendre la bonne décision.

Si la prévision consiste à se baser sur des outils, des données, des éléments, pour anticiper les décisions à prendre afin que celle-ci aient un impact positif dans le futur pour l'entreprise, nous pouvons dire que l'être humain peut aussi se considérer comme un outil, une donnée, un élément, et l'élément phare qui gère tout le processus de décision.

Les résultats obtenus lors de ce stage montrent clairement le potentiel du concept d'un système des prévisions au sein du monde d'entrepreneuriat.

L'approche utilisée dans mon travail est une approche basée sur de nombreuses hypothèses simplificatrices qui sont mentionnées dans la partie Hypothèses, ce qui a rendu les résultats non satisfaisants.

Les modèles utilisés sont des modèles basiques qui donnent en général des résultats satisfaisants, mais le choix des hypothèses a rendu les résultats un peu décalés.

Dans nos jours, les systèmes de prédiction font sujet de nombreuses recherches, et sont des systèmes basés principalement sur des algorithmes du Machine Learning.

Une piste de recherche dans ce sujet est d'implémenter cet algorithme dans un business tout en prenant le contexte en considération et ne pas utiliser les mêmes hypothèses que j'ai déjà admises, cette voie de recherche aboutira sûrement à des meilleurs résultats et une compréhension plus approfondie du problème.

Le domaine de recherche dans ce sujet est arrivé à un stade de maturité qui nécessite un bagage scientifique et technique important, ce qui va me motiver à m'orienter et avoir un parcours

Au niveau personnel, ce stage m'a été très bénéfique. Entre les recherches sur lesquels je me suis lancé, les multiples articles de recherches, j'ai finalement pu avoir du temps pour étudier en profondeur un sujet qui me tient à cœur. Une spécialité qui me passionne. J'ai pu passer au stade d'implémenter mes propres modèles, alors qu'au début du stage je

n'étais qu'un simple enthousiaste.

Ce premier contact m'a permis d'avoir une idée générale sur le monde des entreprises, sur le volet académique j'ai pu tracer mon chemin de mobilité après plusieurs mois d'hésitation.

Bibliographie

- [1] Najib Akesbi, Université citoyenne 2013, HEM Oujda : *“Quelles réformes pour la politique de compensation ?”*
- [2] Ministère des Affaires Générales et de la Gouvernance : *“Réforme de la compensation*
- [3] Nizar Baraka, 2017 : *Rapport du conseil économique, social et environnemental*
- [4] Jacqueline Bouquerel, 2017 : *Le pétrole au Maroc*
- [5] Wiame Idrissi Alami, 2017 : *Les politiques assistancielles de lutte contre la pauvreté : enjeux politiques et effets-retour : étude comparée en France et au Maroc*
- [6] Industrie du futur : l’UTC aux côtés des entreprises : <https://interactions.utc.fr/thematiques/pluridisciplinarite/44-industrie-du-futur-lutc-aux-cotes-des-entreprises/le-data-analytics-au-service-de-lexcellence-industrielle.html>
- [7] *Documents internes, Département Retail, Vivo Energy Maroc*
- [8] Franziska Bell, Slawek Smyl 2018 ,Forecasting at Uber : An Introduction <https://eng.uber.com/forecasting-introduction/>
- [9] Xavier Du Pré, ENSAE Paris 2017 https://www.xavierdupre.fr/app/ensae_teaching_cs/helpsphinx/notebooks/ml_timeseries_base.html
- [10] Prototype de prévision élaboré par l’équipe du Data Science de Facebook <https://facebook.github.io/prophet/>
- [11] Analysez et modélisez des séries temporelles, Open-classrooms : <https://openclassrooms.com/fr/courses/4525371-analysez-et-modelisez-des-series-temporelles/5001311-apprenez-le-lissage-exponentiel-double-et-la-methode-de-holt-winters>

Chapitre 5

Annexe 1 : Description des bases de données utilisées

2016-10-01	99
2016-11-01	132
2016-12-01	191
2017-01-01	132
2017-02-01	165
2017-03-01	196
2017-04-01	165
2017-05-01	196
2017-06-01	196
2017-07-01	265
2017-08-01	356
2017-09-01	224
2017-10-01	228
2017-11-01	231
2017-12-01	228
2018-01-01	231
2018-02-01	198
2018-03-01	261
2018-04-01	277
2018-05-01	297
2018-06-01	309
2018-07-01	380
2018-08-01	413
2018-09-01	281
2018-10-01	277
2018-11-01	297
2018-12-01	297
2019-01-01	330
2019-02-01	264
2019-03-01	284
2019-04-01	343
2019-05-01	264
2019-06-01	297
2019-07-01	376

Freq: MS, Name: Data, dtype: int64

FIGURE 5.1 – Base de données mensuelles

Date	
2019-01-01	363901.92
2019-01-02	338171.22
2019-01-03	364367.40
2019-01-04	374409.70
2019-01-05	390899.76
2019-01-06	311862.98
2019-01-07	317241.86
2019-01-08	335378.34
2019-01-09	365125.96
2019-01-10	432715.38
2019-01-11	366918.92
2019-01-12	317586.66
2019-01-13	346455.04
2019-01-14	323568.94
2019-01-15	315483.38
2019-01-16	414123.78
2019-01-17	393468.39
2019-01-18	392297.43
2019-01-19	429682.05
2019-01-20	327412.47
2019-01-21	360741.78
2019-01-22	389051.46
2019-01-23	410326.77
2019-01-24	428356.11
2019-01-25	398505.24
2019-01-26	414700.65
2019-01-27	377014.68
2019-01-28	338562.42
2019-01-29	371340.69
2019-01-30	363496.98
...	...
2019-06-27	489370.83

FIGURE 5.2 – Base de données journalières 01/01/2019 .. 26/07/2019

Chapitre 6

Annexe 2 : Méthodes classiques

#importation des bibliothèques

```
import pandas as pd
import numpy as np
import sys
import warnings
import itertools
warnings.filterwarnings("ignore")
import statsmodels.api as sm
import statsmodels.tsa.api as smt
import statsmodels.formula.api as smf
from sklearn.metrics import mean_squared_error
import matplotlib.pyplot as plt
get_ipython().run_line_magic('matplotlib', 'inline')
import datetime
import calendar
import seaborn as sns
from statsmodels.tsa.api import ExponentialSmoothing, SimpleExpSmoothing, Holt
```

#Manipulation de la BDD

```
data = pd.read_excel("Produit_2.xlsx")
dates = pd.date_range(start='2016-10-01', freq='MS', periods=len(data))
data.set_index(dates,inplace=True)

sales_ts = data['Data']
```

```
sales_ts

#affichage de la BDD
data = pd.DataFrame(sales_ts.copy())
data.head()

# Splitting Train and Test Length
train_len = int(np.ceil(len(data) * 0.70))

train=data[0:train_len]
test=data[train_len:]
print('Train data length :',len(train))
print('Test data length :',len(test))

x_train = train.drop('Data', axis=1)
x_test = test.drop('Data', axis=1)
y_train = train[['Data']]
y_test = test[['Data']]

#Méthode de la moyenne simple et le tracé

y_hat_avg = test.copy()
y_hat_avg['avg_forecast'] = train['Data'].mean()

plt.figure(figsize=(12,8))
plt.plot(train['Data'], label='Train')
plt.plot(test['Data'], label='Test')
plt.plot(y_hat_avg['avg_forecast'], label='Simple Average Forecast')
plt.legend(loc='best')
plt.show()

# Root Mean Square Error (RMSE)
rmse = np.sqrt(mean_squared_error(test.Data, y_hat_avg.avg_forecast))
rmse = round(rmse, 3)

#Mean Absolute Percentage Error
```

```
abs_error = np.abs(y_test['Data']-y_hat_avg.avg_forecast)
actual = y_test['Data']
mape = np.round(np.mean(abs_error/actual),4)
resultsDf = pd.DataFrame({'Method':['Simple average forecast'], 'RMSE': [rmse]})
```

```
# Storing results in a dataframe
resultsDf = pd.DataFrame({'Method':['Simple average forecast'], 'MAPE': [mape],
'RMSE': [rmse]})
resultsDf = resultsDf[['Method', 'RMSE', 'MAPE']]
```

```
df1 = data.copy()
```

```
#Méthode de la moyenne mobile et le tracé
```

```
#Chercher combien de mois devons nous prendre en consideration pour la méthode
#des moyennes pondérées
# Le choix optimal est prendre les 2 derniers mois pour le calcul (RMSE le plus
#bas)
```

```
for i in range(1,13):
    df1['moving_avg_forecast'] = data['Data'].rolling(i).mean()
    train=df1[0:train_len]
    test=df1[train_len:]
    y_hat_avg = test.copy()
    rmse = np.sqrt(mean_squared_error(test.Data, y_hat_avg.moving_avg_forecast))
    abs_error = np.abs(y_test['Data']-y_hat_avg.moving_avg_forecast)
    actual = y_test['Data']
    mape = np.round(np.mean(abs_error/actual),4)
    print('Rolling window : ',i,'RMSE : ',np.round(rmse,2) , 'MAPE : ',mape)
```

```
# In[209]:
```

```
df1['moving_avg_forecast'] = data['Data'].rolling(2).mean()
```

```
# In[210]:
```

```
#Creating train and test set
```

```
train=df1[0:train_len]
```

```
test=df1[train_len:]
```

```
y_hat_avg = test.copy()
```

```
plt.figure(figsize=(12,6))
```

```
plt.plot(train['Data'], label='Train')
```

```
plt.plot(test['Data'], label='Test')
```

```
plt.plot(y_hat_avg['moving_avg_forecast'], label='Moving Average Forecast')
```

```
plt.legend(loc='best')
```

```
plt.show()
```

```
# Root Mean Square Error (RMSE)
```

```
rmse = np.sqrt(mean_squared_error(test.Data, y_hat_avg.moving_avg_forecast))
```

```
rmse = round(rmse, 3)
```

```
#Mean Absolute Percentage Error
```

```
abs_error = np.abs(y_test['Data']-y_hat_avg.moving_avg_forecast)
```

```
actual = y_test['Data']
```

```
mape = np.round(np.mean(abs_error/actual),4)
```

```
# Storing results in a dataframe
```

```
tempResultsDf = pd.DataFrame({'Method':['Moving average forecast'], 'RMSE':  
[rmse], 'MAPE': [mape] })
```

```
resultsDf = pd.concat([resultsDf, tempResultsDf])
```

```
resultsDf = resultsDf[['Method', 'RMSE', 'MAPE']]
```

```
#Méthode de lissage simple et le tracé
```

```

# create class
model = SimpleExpSmoothing(train['Data'])

#choix du coefficient niveau
#on choisit 0.9 comme valeur optimale
for i in np.arange(0.1,1,0.05):
    model_fit = model.fit(smoothing_level=i,optimized=False)
    y_hat_avg = test.copy()
    y_hat_avg['SES'] = model_fit.forecast(len(test['Data']))
    rmse = np.sqrt(mean_squared_error(test.Data, y_hat_avg.SES))
    rmse = round(rmse, 3)
    abs_error = np.abs(y_test['Data']-y_hat_avg.SES)
    actual = y_test['Data']
    mape = np.round(np.mean(abs_error/actual),4)
    print('smoothing_level : ',np.round(i,1),'RMSE : ' ,np.round(rmse,2) , 'MAPE
        : ',mape)

# fit model
model_fit = model.fit(smoothing_level=0.9,optimized=False)
model_fit.params

y_hat_avg = test.copy()
y_hat_avg['SES'] = model_fit.forecast(len(test['Data']))

# In[219]:

plt.figure(figsize=(12,6))
plt.plot(train['Data'], label='Train')
plt.plot(test['Data'], label='Test')
plt.plot(y_hat_avg['SES'], label='SES')

```

```
plt.legend(loc='best')
```

```
# In[224]:
```

```
# Root Mean Square Error (RMSE)
```

```
rmse = np.sqrt(mean_squared_error(test.Data, y_hat_avg.SES))
```

```
rmse = round(rmse, 3)
```

```
# Mean Absolute Percentage Error
```

```
abs_error = np.abs(y_test['Data']-y_hat_avg.SES)
```

```
actual = y_test['Data']
```

```
mape = np.round(np.mean(abs_error/actual),4)
```

```
# Storing results in a dataframe
```

```
tempResultsDf = pd.DataFrame({'Method':['SES forecast'], 'RMSE': [rmse], 'MAPE':  
[mape] })
```

```
resultsDf = pd.concat([resultsDf, tempResultsDf])
```

```
resultsDf = resultsDf[['Method', 'RMSE', 'MAPE']]
```

```
fit1 = ExponentialSmoothing(np.asarray(train['Data']), seasonal_periods=12 ,
```

```
trend='add', seasonal='None').fit()
```

```
fit1.params
```

```
#Méthode lissage double de HOLT et le tracé
```

```
y_hat_avg['Holt'] = fit1.forecast(len(test))
```

```
# In[227]:
```

```
plt.figure(figsize=(12,8))
```

```
plt.plot( train['Data'], label='Train')
```

```
plt.plot(test['Data'], label='Test')
```

```
plt.plot(y_hat_avg['Holt'], label='Holt')
```



```
plt.legend(loc='best')
plt.show()
```

```
# In[228]:
```

```
# Root Mean Square Error (RMSE)
rmse = np.sqrt(mean_squared_error(test.Data, y_hat_avg.Holt))
rmse = round(rmse, 3)
# Mean Absolute Percentage Error
abs_error = np.abs(y_test['Data']-y_hat_avg.Holt)
actual = y_test['Data']
mape = np.round(np.mean(abs_error/actual),4)
# Storing results in a dataframe
tempResultsDf = pd.DataFrame({'Method':['Holt'], 'RMSE': [rmse], 'MAPE': [mape]})
resultsDf = pd.concat([resultsDf, tempResultsDf])
resultsDf = resultsDf[['Method', 'RMSE', 'MAPE']]
```

```
y_hat_avg = test.copy()
```

```
fit1 = ExponentialSmoothing(np.asarray(train['Data']), seasonal_periods=12 ,
trend='add', seasonal='add').fit()
fit1.params
```

```
#Méthode du lissage triple (Holt winters) et le tracé
```

```
y_hat_avg['Holt_Winter'] = fit1.forecast(len(test))
plt.figure(figsize=(10,6))
plt.plot( train['Data'], label='Train')
plt.plot(test['Data'], label='Test')
plt.plot(y_hat_avg['Holt_Winter'], label='Holt_Winter')
plt.legend(loc='best')
```

```

# Root Mean Square Error (RMSE)
rmse = np.sqrt(mean_squared_error(test.Data, y_hat_avg.Holt_Winter))
rmse = round(rmse, 3)
# Mean Absolute Percentage Error
abs_error = np.abs(y_test['Data']-y_hat_avg.Holt_Winter)
actual = y_test['Data']
mape = np.round(np.mean(abs_error/actual),4)
# Storing results in a dataframe
tempResultsDf = pd.DataFrame({'Method':['Holt Winter forecast - Additive'],
                              'RMSE': [rmse], 'MAPE': [mape] })
resultsDf = pd.concat([resultsDf, tempResultsDf])
resultsDf = resultsDf[['Method', 'RMSE', 'MAPE']]

# In[233]:

y_hat_avg = test.copy()
fit1 = ExponentialSmoothing(np.asarray(train['Data']), seasonal_periods=12,
                             trend='add', seasonal='mul').fit()
fit1.params

# In[234]:

y_hat_avg['Holt_Winter'] = fit1.forecast(len(test))

# In[235]:

plt.figure(figsize=(10,6))
plt.plot( train['Data'], label='Train')
plt.plot(test['Data'], label='Test')
plt.plot(y_hat_avg['Holt_Winter'], label='Holt_Winter')

```

```
plt.legend(loc='best')
```

```
# In[236]:
```

```
# Root Mean Square Error (RMSE)
rmse = np.sqrt(mean_squared_error(test.Data, y_hat_avg.Holt_Winter))
rmse = round(rmse, 3)
# Mean Absolute Percentage Error
abs_error = np.abs(y_test['Data']-y_hat_avg.Holt_Winter)
actual = y_test['Data']
mape = np.round(np.mean(abs_error/actual),4)
# Storing results in a dataframe
tempResultsDf = pd.DataFrame({'Method':['Holt Winter forecast - Multiplicative'],
, 'RMSE': [rmse], 'MAPE': [mape] })
resultsDf = pd.concat([resultsDf, tempResultsDf])
resultsDf = resultsDf[['Method', 'RMSE', 'MAPE']]
resultsDf.index = [i for i in range(len(resultsDf))]
resultsDf
```

Chapitre 7

Annexe 3 : Méthode ARIMA

```
#importation des bibliothèques
import warnings
import itertools
import numpy as np
import matplotlib.pyplot as plt
warnings.filterwarnings("ignore")
plt.style.use('fivethirtyeight')
import pandas as pd
import statsmodels.api as sm
import matplotlib

matplotlib.rcParams['axes.labelsize'] = 14
matplotlib.rcParams['xtick.labelsize'] = 12
matplotlib.rcParams['ytick.labelsize'] = 12
matplotlib.rcParams['text.color'] = 'k'

#Manipulation de la BDD
df = pd.read_excel("SalesAljazira.xlsx")
#affichage de la BDD
df.head()

df = df.groupby('Date')['Data'].sum().reset_index()
df = df.set_index('Date')
df.index

#Description de la BDD
df.describe()
```

```
#Tracé de la BDD
```

```
df.plot(figsize=(15, 6))
plt.show()
```

```
# Décomposition de la BDD
```

```
from pylab import rcParams
rcParams['figure.figsize'] = 18, 8
decomposition = sm.tsa.seasonal_decompose(df, model='additive')
fig = decomposition.plot()
plt.show()
```

```
#Initialisation de la méthode
```

```
p = d = q = range(0, 2)
pdq = list(itertools.product(p, d, q))
seasonal_pdq = [(x[0], x[1], x[2], 12) for x in list(itertools.product(p, d, q))]
print('Examples of parameter combinations for Seasonal ARIMA...')
print('SARIMAX: {} x {}'.format(pdq[1], seasonal_pdq[1]))
print('SARIMAX: {} x {}'.format(pdq[1], seasonal_pdq[2]))
print('SARIMAX: {} x {}'.format(pdq[2], seasonal_pdq[3]))
print('SARIMAX: {} x {}'.format(pdq[2], seasonal_pdq[4]))

for param in pdq:
    for param_seasonal in seasonal_pdq:
        try:
            mod = sm.tsa.statespace.SARIMAX(df,
                                              order=param,
                                              seasonal_order=param_seasonal,
                                              enforce_stationarity=False,
                                              enforce_invertibility=False)

            results = mod.fit()
            print('ARIMA{}x{}12 - AIC:{}'.format(param, param_seasonal, results.aic))
        except:
            continue
```

```
mod = sm.tsa.statespace.SARIMAX(df,
                                order=(1, 1, 1),
                                seasonal_order=(1, 1, 1, 12),
                                enforce_stationarity=False,
                                enforce_invertibility=False)

results = mod.fit()
print(results.summary().tables[1])
```

```
results.plot_diagnostics(figsize=(16, 8))
plt.show()
```

#Prévision

```
pred = results.get_prediction(start=pd.to_datetime('2019-05-01'), dynamic=False)
pred_ci = pred.conf_int()
ax = df['2015:'].plot(label='observed')
pred.predicted_mean.plot(ax=ax, label='One-step ahead Forecast', alpha=.7, figsize=(16, 8))
ax.fill_between(pred_ci.index,
                pred_ci.iloc[:, 0],
                pred_ci.iloc[:, 1], color='k', alpha=.2)
ax.set_xlabel('Date')
ax.set_ylabel('Sales')
plt.legend()
plt.show()
```

```
y_forecasted = pred.predicted_mean
y_t = df['2019-05-01:']
```

```
y_truth=y_t.squeeze()
```

y_truth

```
mse = ((y_forecasted - y_truth) ** 2).mean()
print('The Mean Squared Error of our forecasts is {}'.format(round(mse, 2)))
```

In[21]:

```
print('The Root Mean Squared Error of our forecasts is {}'.format(round(np.sqrt(mse),
```

Choix du nombre des mois

```
pred_uc = results.get_forecast(steps=100)
pred_ci = pred_uc.conf_int()
ax = df.plot(label='observed', figsize=(14, 7))
pred_uc.predicted_mean.plot(ax=ax, label='Forecast')
ax.fill_between(pred_ci.index,
                pred_ci.iloc[:, 0],
                pred_ci.iloc[:, 1], color='k', alpha=.25)
ax.set_xlabel('Date')
ax.set_ylabel(' Sales')
plt.legend()
plt.show()
```

#Enregistrement du fichier Excel

```
pred_uc.predicted_mean.to_excel("Results__Daily.xlsx")
```