**Analyse des Disparités Régionales et Prévision des Prix du Gasoil Non Routier en France**

**1. Introduction**

Dans un contexte économique en constante évolution, le marché du Gasoil Non Routier (GNR) en France joue un rôle crucial pour de nombreux secteurs, notamment l'agriculture, la sylviculture et les travaux publics. Les fluctuations de prix de ce carburant essentiel ont des répercussions directes sur la compétitivité des entreprises et l'économie régionale. Notre étude se propose d'explorer en profondeur les dynamiques complexes qui régissent ce marché, en mettant l'accent sur les disparités régionales et les facteurs influençant les variations de prix.

L'introduction du GNR en 2011, dans le cadre des efforts de réduction des émissions polluantes, a marqué un tournant dans le paysage énergétique français. Depuis lors, les politiques énergétiques et fiscales ont continué d'évoluer, comme en témoigne la récente décision gouvernementale de janvier 2024 d'annuler la hausse progressive de la taxe sur le GNR prévue jusqu'en 2030. Cette mesure, visant à alléger la charge fiscale des agriculteurs, pourrait avoir des implications significatives sur la demande et, par conséquent, sur les prix du GNR.

Dans ce contexte, notre étude vise à répondre à une question centrale : Quels sont les facteurs déterminants des disparités régionales des prix du GNR en France, et comment peut-on prédire ces variations de prix de manière fiable ? Pour y répondre, nous nous appuyons sur des techniques avancées de data science et d'intelligence artificielle, ouvrant ainsi de nouvelles perspectives dans l'analyse de ce marché complexe.

**2. Présentation des Données**

**2.1 Collecte et Sources des Données**

Notre étude s'appuie sur un ensemble de données riche et diversifié, collecté quotidiennement depuis avril 2024. La méthode principale de collecte a été le web scraping, une technique permettant d'extraire automatiquement des informations de sites web. Cette approche nous a permis d'obtenir des données actualisées et détaillées sur les prix du GNR auprès de multiples distributeurs dans plus de 1900 codes postaux uniques en France.

En complément de ces données de prix, nous avons intégré des informations provenant de diverses sources :

1. Données économiques : L'indice CNR (Comité National Routier), les taux de change USD/EUR, et le cours du baril de pétrole en USD ont été obtenus à partir de sources financières et économiques fiables.
2. Données démographiques et géographiques : Les informations sur la population, la densité de population, la latitude et la longitude ont été extraites de bases de données gouvernementales.
3. Données météorologiques : Les températures moyennes ont été collectées auprès de services météorologiques nationaux.
4. Données sur l'activité économique locale : Le nombre d'entreprises par secteur d'activité a été obtenu à partir de registres du commerce et de l'industrie.

**2.2 Structure de la Base de Données**

Les données collectées ont été structurées dans une base de données relationnelle, hébergée sur Google Cloud Platform (GCP) via BigQuery. Cette structure nous permet une gestion efficace et des analyses rapides de grands volumes de données. Voici un aperçu des principales variables de notre base de données :

* Variables temporelles : date, année-mois
* Variables géographiques : code postal, département, région, latitude, longitude
* Variables démographiques : population, surface, tranche\_population, densité\_population
* Variables économiques : indice\_CNR, USD\_to\_EUR, EUR\_to\_USD, cours\_baril\_en\_USD
* Variables météorologiques : tmoy (température moyenne)
* Variables d'activité économique : nb\_entreprise\_ensemble, nb\_entreprise\_industrie, nb\_entreprise\_construction, etc.
* Variable cible : prix du GNR

Cette structure de données nous offre une granularité fine tant au niveau géographique que temporel, permettant des analyses détaillées des variations de prix du GNR.

**3. Méthodologie et Analyse**

Notre approche méthodologique combine des techniques statistiques traditionnelles avec des méthodes avancées de machine learning et d'analyse géospatiale. Cette synergie nous permet d'explorer les données sous différents angles et de capturer des patterns complexes dans les variations de prix du GNR.

**3.1 Distribution Géographique des Prix**

Pour visualiser les disparités régionales, nous avons créé une carte choroplèthe de la France (Figure 1) représentant les prix moyens du GNR par région. Cette visualisation révèle des contrastes saisissants :

Show Image

*Figure 1: Distribution géographique des prix moyens du GNR*

L'analyse de cette carte met en lumière des disparités régionales significatives :

* Les régions du Nord-Est et du Sud-Est affichent des prix systématiquement plus élevés (>1350), une tendance que nous attribuons à leur éloignement des principales sources d'approvisionnement et aux coûts de transport associés.
* En contraste, les régions du Centre et de l'Ouest bénéficient de prix plus modérés (1250-1300), suggérant une meilleure accessibilité logistique.

Ces observations soulignent l'importance cruciale des facteurs géographiques et logistiques dans la formation des prix du GNR, un aspect que les politiques de tarification et d'approvisionnement devraient prendre en compte.

**3.2 Évolution Temporelle des Prix**

Pour comprendre les dynamiques temporelles, nous avons analysé l'évolution des prix par région sur la période étudiée (Figure 2).

Show Image

*Figure 2: Évolution du prix du GNR par région*

Ce graphique révèle plusieurs tendances clés :

1. Une tendance générale à la baisse des prix de mai à octobre 2024, probablement influencée par des facteurs macroéconomiques et saisonniers.
2. Une période de forte volatilité en juin-juillet, suivie d'une relative stabilisation, suggérant des perturbations temporaires du marché.
3. Une convergence progressive des prix entre les régions vers la fin de la période, indiquant potentiellement un rééquilibrage du marché.

Particulièrement notable est le comportement atypique de certaines régions :

* La Normandie se distingue par des pics de prix extrêmes, notamment début juin, nécessitant une investigation approfondie des facteurs locaux spécifiques.
* L'Occitanie montre une volatilité persistante, suggérant une sensibilité accrue aux fluctuations du marché.
* L'Île-de-France maintient une relative stabilité des prix, généralement inférieurs à la moyenne nationale, possiblement due à une demande stable et une infrastructure logistique efficace.

**3.3 Analyse Nationale et Distribution Globale des Prix**

Pour compléter notre analyse régionale, nous avons examiné l'évolution du prix moyen national (Figure 3) et la distribution globale des prix (Figure 4).

Show Image

*Figure 3: Évolution du prix moyen du GNR*

La courbe nationale révèle :

1. Un pic initial en mai-juin 2024
2. Un plateau élevé en juillet
3. Une baisse progressive d'août à septembre
4. Une légère remontée fin octobre

Cette évolution suggère des perturbations importantes du marché en début de période, suivies d'un ajustement progressif et d'une possible influence saisonnière sur les prix.

Show Image

*Figure 4: Distribution des prix du GNR*

L'histogramme de distribution des prix montre :

* Une distribution approximativement normale avec une légère asymétrie positive
* Un pic principal entre 1250 et 1300
* Une queue de distribution étendue vers les prix élevés

Ces observations indiquent que, bien que la majorité des prix se concentre autour d'une valeur moyenne, les hausses de prix extrêmes sont plus fréquentes que les baisses extrêmes, soulignant la volatilité potentielle du marché du GNR.

**3.4 Analyse des Corrélations**

Pour identifier les facteurs influençant les prix du GNR, nous avons réalisé une analyse de corrélation approfondie (Figure 5).

Show Image

*Figure 5: Matrice de corrélation des variables clés*

Les résultats les plus significatifs montrent :

* Une forte corrélation positive entre les prix du GNR et l'indice CNR (0.617) ainsi qu'avec le cours du baril en USD (0.612), soulignant l'importance des facteurs macroéconomiques.
* Des corrélations opposées avec les taux de change (USD\_to\_EUR : -0.604, EUR\_to\_USD : 0.599), reflétant l'impact des fluctuations monétaires sur les prix.
* Des corrélations négatives modérées avec le nombre d'entreprises et la population, suggérant une tendance à des prix légèrement plus bas dans les régions plus peuplées et économiquement actives.

Ces corrélations mettent en lumière la complexité des facteurs influençant les prix du GNR et soulignent l'importance d'une approche multifactorielle dans l'analyse et la prédiction des prix.

**4. Discussion et Implications**

Notre analyse révèle un marché du GNR caractérisé par une forte hétérogénéité spatiale et temporelle. Les facteurs macroéconomiques, tels que l'indice CNR et le cours du pétrole, apparaissent comme les principaux déterminants des prix, tandis que l'influence des facteurs géographiques, bien que présente, semble moins prononcée qu'initialement supposé.

Les disparités régionales observées soulèvent des questions importantes sur l'équité d'accès au GNR et ses implications pour la compétitivité des entreprises dans différentes régions. Les régions confrontées à des prix systématiquement plus élevés, comme le Nord-Est et le Sud-Est, pourraient bénéficier de politiques ciblées visant à atténuer ces disparités, par exemple à travers des investissements dans l'infrastructure logistique ou des mesures fiscales adaptées.

La volatilité des prix, particulièrement prononcée dans certaines régions comme la Normandie, appelle à une vigilance accrue et à la mise en place de mécanismes de stabilisation. Les entreprises opérant dans ces régions pourraient envisager des stratégies de couverture ou de diversification pour atténuer les risques liés aux fluctuations de prix.

**5. Conclusion et Perspectives**

Cette étude offre une compréhension approfondie des dynamiques de prix du GNR en France, révélant la complexité et la variabilité du marché. Les insights obtenus fournissent une base solide pour la prise de décision stratégique, tant pour les acteurs publics que privés du secteur.

Pour approfondir cette analyse, plusieurs pistes de recherche future se dessinent :

1. Extension de l'étude sur une période plus longue pour identifier les tendances à long terme et les cycles saisonniers.
2. Intégration de variables supplémentaires, telles que les politiques régionales spécifiques ou les infrastructures de transport, pour affiner notre compréhension des facteurs influençant les prix.
3. Développement de modèles prédictifs plus sophistiqués, intégrant des techniques de deep learning pour capturer des patterns complexes dans les variations de prix.
4. Réalisation d'une analyse comparative à l'échelle européenne pour contextualiser les spécificités du marché français et anticiper l'impact des politiques énergétiques sur la stabilité des prix du GNR.

En conclusion, cette étude constitue une base solide pour une gestion plus éclairée et stratégique du marché du GNR en France, ouvrant la voie à des politiques plus adaptées et à une meilleure anticipation des fluctuations de prix dans ce secteur crucial de l'économie.