Une image contenant Graphique, logo, Police, rouge

Description générée automatiquement

# MTI820-01- Entrepôts de données et intelligence d'affaires.

Analyse de l’impact du sport et de l’environnement sur la santé mentale et physique.

Ridah LARBI-DAOUADJI, Fanny SALMON, Emma CARILLO

Session : Hiver 2025

Terminé le 2025

**I. Introduction**

A) Problématique

B) Contexte et enjeux

C) Objectifs du projet

**II. Entrée en matière**

A) Impact du sport sur la santé mentale et physique

B) Influence des conditions environnementales sur l’activité physique

C) Corrélation entre pollution, climat et bien-être mental

D) Études et recherches antérieures

**III. Méthodologie**

A) Sources de données utilisées

1) Données environnementales (qualité de l’air, météo)

2) Données de santé mentale et physique

3) Données sportives et activité physique

B) Stratégie d’analyse et modélisation

1) Techniques de Machine Learning

2) Algorithmes de corrélation

C) Visualisation et Business Intelligence

1) Utilisation de Power BI / Tableau

2) Indicateurs clés

D) Recommandations et scénarios prédictifs

1) Analyse des tendances

2) Système d’alerte

**IV. Planification et réalisation du projet**

A) Période de travail et calendrier

B) Répartition des tâches

C) Outils et technologies utilisées

D) Difficultés et solutions envisagées

**V. Résultats et analyse des données**

A) Corrélation entre sport et santé mentale

B) Impact des conditions météorologiques sur la motivation et la pratique sportive

C) Influence de la pollution sur la performance physique

D) Comparaison entre différentes zones géographiques

**VI. Discussion et recommandations**

A) Interprétation des résultats

B) Limites de l’étude et biais possibles

C) Propositions d’amélioration et perspectives

**VII. Conclusion**

A) Synthèse des résultats

B) Implications pour la santé publique

C) Pistes pour de futures recherches

**VIII. Références**

Liste des articles et sources citées

**IX. Annexes**

A) Visualisations

B) Documentation technique

1. Introduction:
2. Problématique:

La santé mentale et physique est influencée par plusieurs facteurs, notamment la pratique sportive et les conditions environnementales. L’Organisation mondiale de la santé (OMS) recommande au moins 150 minutes d’exercice modéré par semaine, permettant de réduire d’environ 30 % les risques de maladies cardiovasculaires, de diabète et d’autres affections chroniques [1]. Toutefois, ces bienfaits peuvent être atténués par les contraintes environnementales qui influencent la pratique du sport et ses effets sur la santé.

En milieu urbain, la pollution atmosphérique est un facteur de risque majeur. En 2019, l’OMS estimait que 68 % des décès prématurés liés à la pollution extérieure étaient dus à des cardiopathies ischémiques et des accidents vasculaires cérébraux. De plus, 14 % étaient attribués à des bronchopneumopathies chroniques obstructives, 14 % à des infections aiguës des voies respiratoires inférieures et 4 % à des cancers du poumon [2]. L’exposition prolongée aux particules fines (PM2.5) entraîne des troubles respiratoires, une réduction des capacités cardiovasculaires et un affaiblissement du système immunitaire, augmentant ainsi le risque de maladies chroniques.

Les effets de la pollution ne sont pas seulement physiques. Une exposition prolongée à une mauvaise qualité de l’air peut accroître le stress oxydatif et l’inflammation neurologique, augmentant les risques d’anxiété, de dépression et de troubles cognitifs. Elle affecte aussi la qualité du sommeil et accroît la fatigue, influençant négativement le bien-être mental et la productivité [3].

Les conditions météorologiques jouent aussi un rôle clé. Une étude de Tucker et Gilliland (2007) montre que les températures extrêmes et les précipitations réduisent la pratique d’activités physiques en extérieur [4]. Ces conditions peuvent aussi provoquer des problèmes de santé, comme la déshydratation ou l’hypothermie. Sur le plan mental, elles influencent l’humeur et la motivation, accentuant le stress et diminuant le bien-être global.

L’accès aux espaces verts est un facteur protecteur. Mitchell et Popham (2008) ont constaté que les personnes vivant à proximité de parcs ou de forêts présentaient des niveaux de stress réduits de 20 % par rapport aux habitants de zones densément urbanisées [5].

Comprendre l’impact des conditions environnementales (pollution, météo, accès aux espaces verts) sur la pratique du sport permettrait d’optimiser les recommandations en activité physique et d’adapter les politiques publiques pour favoriser une pratique sportive bénéfique, notamment en milieu urbain.

1. Contexte et enjeux :

La santé publique est un enjeu majeur, particulièrement dans les sociétés urbanisées où les conditions environnementales peuvent aggraver les risques de maladies physiques et mentales. À travers ce projet, nous cherchons à mieux comprendre ces influences afin de proposer des outils concrets d’aide à la décision. Grâce à l’analyse de données québécoises, nous mettrons en lumière les effets de la pollution et des conditions climatiques sur la motivation à pratiquer une activité physique et sur l’état mental des individus.  
Ce projet s’inscrit dans une démarche de prévention et d’optimisation des politiques de santé publique, en exploitant les outils modernes de l’intelligence d’affaires.

C) Objectifs du projet :

Notre projet vise à développer une solution de Business Intelligence, permettant de visualiser l’impact des conditions environnementales sur l’activité physique et la santé mentale des Québécois. L’interface offrira des visualisations dynamiques et des analyses avancées à partir de données environnementales et sportives.

1. ***Développement d’un tableau de bord interactif :***

* Crée un tableau de bord regroupant les indicateurs clés liés à la pratique sportive et aux conditions environnementales et à la santé mentale.
* Intégration des données sur la qualité de l’air et observer leur corrélation avec la fréquence et l’intensité des activités physiques, ainsi que leur impact sur le bien être mental et physique.
* Suivre les tendances de température, humidité et précipitations pour analyser leur impact sur la motivation et l’humeur de la population.
* Visualisation de la proximité des espaces verts et évaluer leur influence sur le bien être mentale et physique, notamment en lien avec la réduction du stress et l’amélioration de la santé cognitive.

1. ***Analyse et modélisation des données :***

* Associer des indicateurs comme durée, fréquence, intensité de l’exercice avec les niveaux de stress et d’anxiété pour montrer la corrélation entre activité physique et bien être mental.
* Déterminer à partir des données environnementales à quel niveau de pollution ou température extrême deviennent négatif, tant sur la santé physique que mentale (fatigue, stress, troubles respiratoires).
* Utilisation de Machine Learning pour intégrer des algorithmes prédictifs permettant d’anticiper les jours optimaux pour pratiquer du sport en fonction des conditions environnementales, en évaluant leur impact global sur la santé, notamment la performance physique (énergie, endurance, récupération) et le bien-être mental (réduction du stress, amélioration de l’humeur)

1. ***Génération de recommandations :***

* Intégrer des recommandations personnalisées basées sur l’analyse des données
* Mettre en avant les périodes faiblement polluées et climatiquement favorables
* Proposer des améliorations pour l’urbanisme
* Mettre en place un système d’alertes pour informer les utilisateurs sur les jours conseillées ou déconseilles à la pratique sportive à cause de la pollution ou de la météo.

1. Entrée en matière :
   1. Impact du sport sur la santé mentale et physique :

La pratique régulière d’une activité physique est reconnue pour ses nombreux bienfaits sur la santé physique, mais également sur la santé mentale. Elle permet de renforcer le système cardiovasculaire, d’améliorer la capacité respiratoire, de réguler le poids corporel et de prévenir certaines maladies chroniques telles que le diabète de type 2 ou l’hypertension [1].

Sur le plan psychologique, l’exercice stimule la production d’endorphines, de sérotonine et de dopamine – des neurotransmetteurs liés au bien-être et à la régulation de l’humeur. Des études démontrent que l’activité physique contribue à réduire les symptômes de stress, d’anxiété et de dépression [2]. Elle améliore également la qualité du sommeil, la concentration et la confiance en soi [3]. Au Québec, les politiques de santé publique encouragent la pratique sportive dans une perspective de prévention et de promotion de la santé globale [4].

* 1. Influence des conditions environnementales sur l’activité physique :

L’environnement joue un rôle central dans la fréquence, l’intensité et la qualité de la pratique d’une activité physique. Des températures extrêmes, une forte humidité, des précipitations importantes ou une mauvaise qualité de l’air peuvent décourager les individus à sortir ou à pratiquer un sport. Ces contraintes sont particulièrement visibles en milieu urbain, où les épisodes de chaleur ou de froid intenses, combinés à la pollution, limitent l’accès à une activité physique sécuritaire [5].

En revanche, des conditions favorables comme une météo clémente, des infrastructures sportives accessibles ou la proximité d’espaces verts favorisent une plus grande assiduité dans la pratique sportive. Comprendre ces variables environnementales est donc crucial pour encourager une activité physique régulière [6]

* 1. Corrélation entre pollution, climat et bien-être mental :

Les impacts de la pollution atmosphérique et du climat sur le bien-être mental sont de plus en plus étudiés. Une exposition prolongée à des polluants comme les particules fines (PM2.5), le dioxyde d’azote (NO2) ou l’ozone (O3) est associée à un stress oxydatif accru, à des inflammations cérébrales et à une augmentation des troubles anxieux et dépressifs [7].

De même, la météo peut affecter l’humeur : les journées grises, humides ou très froides sont souvent liées à une baisse de motivation, une fatigue accrue et une diminution du moral général [8].

Ainsi, une mauvaise qualité de l’air ou des conditions météorologiques défavorables peuvent avoir un double effet négatif : limiter la pratique sportive et détériorer directement la santé mentale. C’est ce lien complexe que notre projet vise à explorer à travers les données.

* 1. Études et recherches antérieures

Plusieurs études internationales et québécoises soutiennent l’existence de corrélations entre activité physique, santé mentale et conditions environnementales. L’Organisation mondiale de la santé (OMS) a publié de nombreux rapports sur les effets de la pollution sur la santé physique et mentale [1][7].

Des travaux comme ceux de **Tucker et Gilliland (2007)** ont montré que les conditions météorologiques extrêmes réduisent significativement les niveaux d’activité physique [5]. D’autres recherches soulignent l’effet protecteur des espaces verts sur le stress et la dépression (**Mitchell et Popham, 2008**) [9].

Au Québec, l’Institut national de santé publique (INSPQ) et **Statistique Canada** publient régulièrement des données sur les habitudes de vie, la santé mentale et les conditions environnementales, qui servent de base empirique à notre projet [10]. Ces ressources permettent une analyse contextualisée et adaptée aux réalités de la population québécoise.

1. Méthodologie :
   1. Sources de données utilisées :
2. Données environnementales (qualité de l’air, méteo) :

Pour mesurer l’impact de l’environnement sur la santé mentale et la pratique sportive, nous avons intégré deux types de données environnementales :

Les données de qualité de l’air ont été extraites du fichier rsqa-indice-qualite-air-2022-2024.csv, issu du Réseau de surveillance de la qualité de l’air (RSQA). Ce jeu de données comprend des mesures horaires de plusieurs polluants atmosphériques, tels que les particules fines (PM2.5), le dioxyde d’azote (NO₂), l’ozone (O₃), le monoxyde de carbone (CO) et le dioxyde de soufre (SO₂). Après un filtrage par identifiant de station (stationId), seules les stations localisées à Montréal ont été conservées. Les données ont ensuite été agrégées pour obtenir des moyennes mensuelles par polluant.

Les données météorologiques, quant à elles, ont été collectées à partir des fichiers mensuels fournis par Environnement Canada. Ces fichiers contiennent des variables telles que la température moyenne (Temp (°C)), l’humidité relative (Hum. rel (%)), la vitesse du vent, le refroidissement éolien, les précipitations et le point de rosée. Les douze fichiers .csv (un par mois) ont été nettoyés, harmonisés puis concaténés, avant d’être moyennés par mois afin d’obtenir un profil météorologique mensuel cohérent.

1. Données de santé mentale et physique :

Ces données ont été simulées à partir de sources publiques (Statistique Canada, INSPQ) en représentant :

* Le **niveau de stress dominant** (qualitatif : "faible", "modéré", "élevé")
* Le **niveau de bien-être dominant** (qualitatif : "faible", "moyen", "bon")  
  Ces catégories ont ensuite été **encodées en scores numériques** (stress\_score et bien\_etre\_score) pour faciliter l’analyse corrélationnelle.

1. Données sportives :

Le fichier Activit\_\_physique\_mensuelle\_simul\_e\_-\_Qu\_bec\_2023.csv contient des données simulées sur l’activité physique des Québécois, incluant :

* Le **nombre moyen de sportifs par jour**
* La **durée moyenne des activités** (en minutes)
* Le **type d’activité dominante** par mois  
  Ce fichier a été nettoyé et fusionné avec les autres sources via une colonne commune : le mois (mois).
  1. Stratégie d’analyse et modélisation :

1. Nettoyage des données :

Les fichiers météorologiques, fournis sous forme de fichiers .csv mensuels par Environnement Canada, ont d’abord été nettoyés. Ce nettoyage a consisté à supprimer les colonnes entièrement vides ou inutiles (indicateurs, symboles, flags), puis à imputer les valeurs manquantes dans les colonnes numériques par leur moyenne mensuelle. Une fois ces opérations effectuées, les fichiers ont été sauvegardés dans un dossier spécifique (METEO\_NETTOYE) pour garantir leur qualité avant agrégation.

Concernant les données de qualité de l’air, issues du Réseau de surveillance de la qualité de l’air (RSQA), nous avons filtré uniquement les enregistrements de l’année 2023. La colonne "heure", considérée comme non pertinente pour une agrégation mensuelle, a été supprimée. Nous avons également retiré les lignes incomplètes, notamment celles avec des valeurs manquantes dans les colonnes critiques comme stationId, polluant, valeur ou date. Enfin, un nouveau fichier propre (rsqa\_air\_montreal\_2023\_nettoye.csv) a été généré pour permettre une agrégation par mois et par polluant

1. Analyse corrélationnelle :

Afin de comprendre les relations entre les différentes variables environnementales, sportives et psychologiques, nous avons réalisé une analyse corrélationnelle à l’aide du langage Python. Les bibliothèques **pandas**, **seaborn** et **matplotlib** ont été utilisées pour traiter les données, calculer les corrélations et visualiser les résultats.

Le processus a commencé par l’importation des données fusionnées (météo, pollution, activité physique, santé mentale), nettoyées et harmonisées. Nous avons ensuite sélectionné un ensemble de variables quantitatives pertinentes telles que la température, l’humidité relative, le nombre moyen de sportifs par jour, les scores de bien-être, ainsi que les niveaux de polluants (PM2.5, NO2, O3, CO, SO2).

Une **matrice de corrélation de Pearson** a été générée pour mesurer la force et la direction des relations linéaires entre ces variables. La **corrélation de Pearson (ρ)** prend une valeur comprise entre -1 (corrélation négative parfaite) et +1 (corrélation positive parfaite). Une valeur proche de 0 indique une absence de lien linéaire.

Nous avons ensuite extrait uniquement les corrélations considérées comme **fortes** (|ρ| ≥ 0.5), en éliminant les doublons et les auto-corrélations. Ces résultats ont été affichés dans la console pour une lecture textuelle, mais aussi sous forme de **carte thermique (heatmap)**, facilitant leur interprétation visuelle.

Ce processus nous a permis d’identifier des **interactions significatives** entre, par exemple, le niveau de pollution et le bien-être, ou encore entre les conditions météorologiques et la fréquence de pratique sportive. Ces corrélations guideront nos visualisations dans Power BI et nourriront nos recommandations stratégiques en matière de santé publique.

## V) Références

[1] Organisation mondiale de la santé. (2022). Activité physique.  
<https://www.who.int/fr/news-room/fact-sheets/detail/physical-activity>

[2] Organisation mondiale de la santé. (2023). Qualité de l’air ambiant et santé.  
<https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health>

[3] Tucker, P., & Gilliland, J. (2007). The effect of season and weather on physical activity: A systematic review. Public Health, 121(12), 909–922.  
<https://www.sciencedirect.com/science/article/abs/pii/S0033350607001394>

[4] Mitchell, R., & Popham, F. (2008). Effect of exposure to natural environment on health inequalities: an observational population study. The Lancet, 372(9650), 1655–1660.  
<https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(08)61689-X>

[5] Canadian Psychological Association. (2016). L’activité physique, la santé mentale et la motivation.  
<https://cpa.ca/docs/File/Publications/FactSheets/FS_Lactivite_physique_la_sante_mentale_la_motivation_FR_novembre2016.pdf>

[6] Peluso, M. A., & Guerra de Andrade, L. H. (2005). Physical activity and mental health: the association between exercise and mood. Clinics, 60(1), 61–70. <https://www.sciencedirect.com/science/article/pii/S1807593222032173?via%3Dihub>

**[7]** Gouvernement du Québec. (2023). *Politique gouvernementale de prévention en santé*.  
 <https://publications.msss.gouv.qc.ca/msss/document-003025/> .

**[8]** Frumkin, H. (2001). Beyond toxicity: Human health and the natural environment. American Journal of Preventive Medicine, 20(3), 234–240. <https://doi.org/10.1016/S0749-3797(00)00317-2>

**[9]** Obradovich, N., et al. (2018). Climate change may affect mental health. Proceedings of the National Academy of Sciences (PNAS), 115(43), 10953–10958. <https://doi.org/10.1073/pnas.1801528115>

[10] Institut national de santé publique du Québec (INSPQ). Enquête québécoise sur la santé de la population.  
<https://www.inspq.qc.ca>