

Alexandre MAURICE  
Guillaume TRITSCH  
Nicolas DARGAZANLI



## S.A.É. 2.04 : Exploitation d'une base de données

### Étape 2 : Production statistique

Compte-rendu

Analyse concernant les accidents en fonction de jours particuliers (vacances, ponts, jours fériés...)

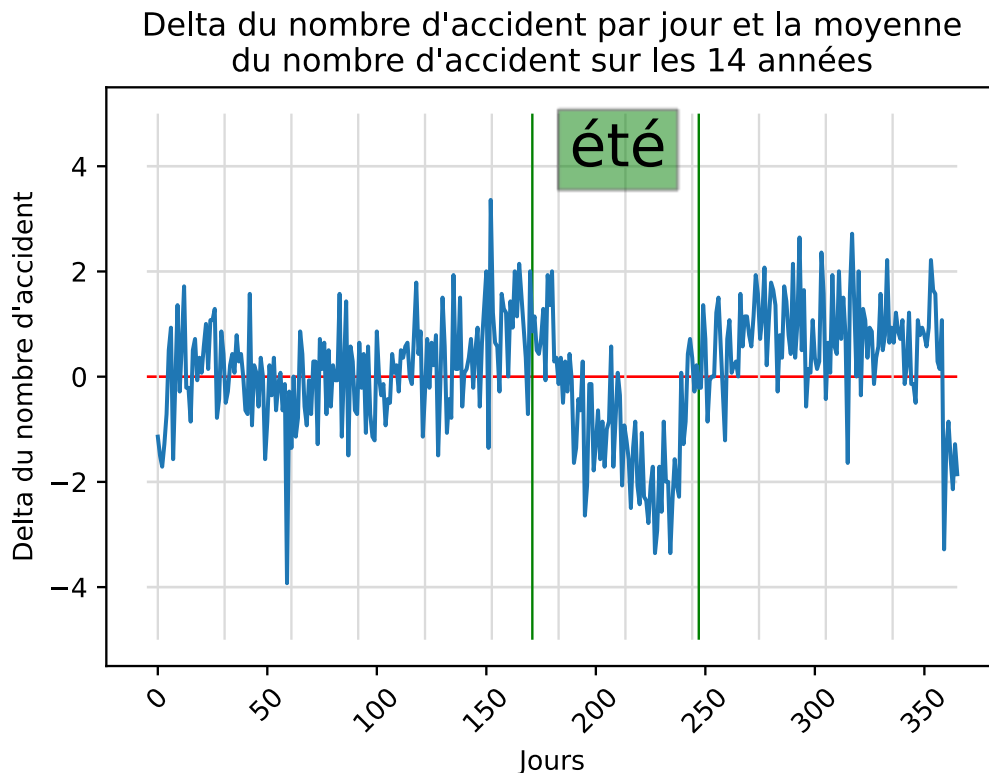
## Notre démarche :

### Question 1 :

*Afficher sur une année la moyenne du nombre d'accidents par jours. Le comparer avec la moyenne du nombre d'accidents sur toute la période entre 1984 et 1998. Un  $\Delta$  peut être utilisé pour mieux mettre en valeur ces écarts.*

Nous voulions montrer ici une évolution qui aurait pu être intéressante en fonction des périodes : en effet, nous soupçonnions qu'il y aurait moins d'accidents pendant les périodes de vacances ou de fêtes, comme l'été ou la fin de l'année, car en grande majorité les accidents sont produits sur un trajet habituel.

Il nous fallait alors avoir un nombre de référence pour notre  $\Delta$ , et nous avons choisi la moyenne du nombre d'accidents par jour (peu importe le jour). Ensuite, nous affichons l'évolution de la différence entre le référentiel et le nombre d'accidents par jour sur l'année.

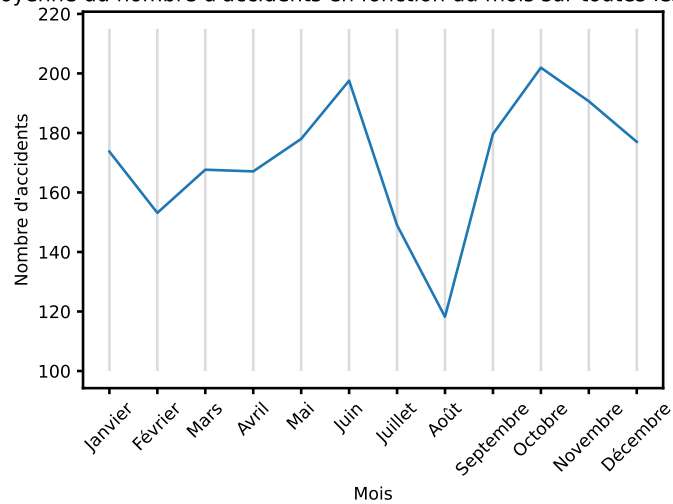


Nous obtenons alors ce graphe, où nous pouvons noter deux grandes périodes ayant connu moins d'accidents : l'été (début de l'été, jusqu'à la rentrée scolaire) et la fin de l'année (qui est malheureusement coupée en deux...), confirmant notre hypothèse de départ.

Nous avons décidé de conserver ce graphe car il démontrait bien l'incidence des périodes particulières par rapport au à la fréquence des accidents.

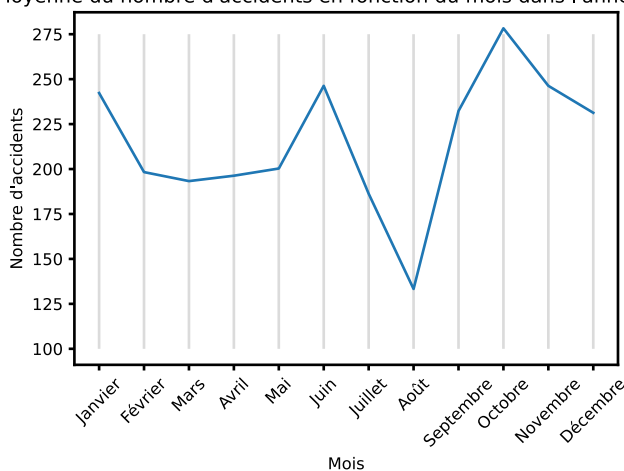
Nous avons aussi réalisé un graphe similaire mais sans tenir compte du  $\Delta$ , avec le nombre d'accidents par mois, mais nous ne l'avons pas gardé car il était moins pertinent vis-à-vis des deux autres graphes pour une année particulière (1984 et 1996).

Moyenne du nombre d'accidents en fonction du mois sur toutes les années

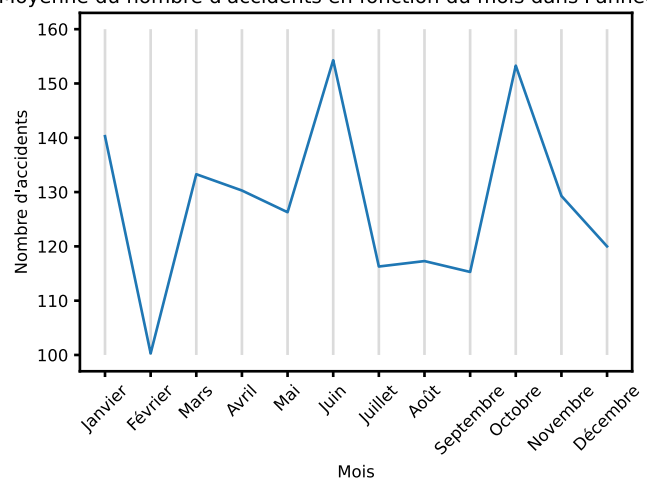


Les deux graphes conservés sont donc ceux-ci :

Moyenne du nombre d'accidents en fonction du mois dans l'année 1984



Moyenne du nombre d'accidents en fonction du mois dans l'année 1996



Question 2 :

*Étudier les types de parties impliquées lors d'un accident des dates particulières, et le comparer avec des jours dits « normaux » ou peu influencés. On pourrait éventuellement étudier aussi la cause de ces accidents. S'intéresser à une période succédant un jour particulier peut être intéressant.*

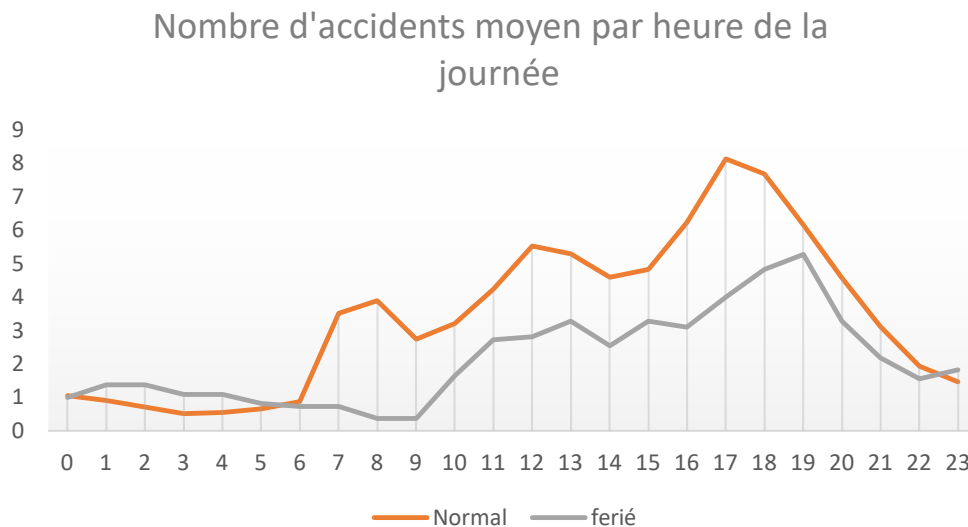
Nous avons décidé de ne pas réaliser cette requête car les résultats n'étaient pas satisfaisants à cause d'un manque de lisibilité, d'un graphique compliqué à interpréter, et de résultats peu significatifs. Malheureusement, nous ne disposons plus des graphes car ils se sont fait effacer dans le cache de Spyder.

Question 3 :

*Comparer les heures « les plus à risque » de produire un accident un jour férié et les comparer avec les heures les plus risquées.*

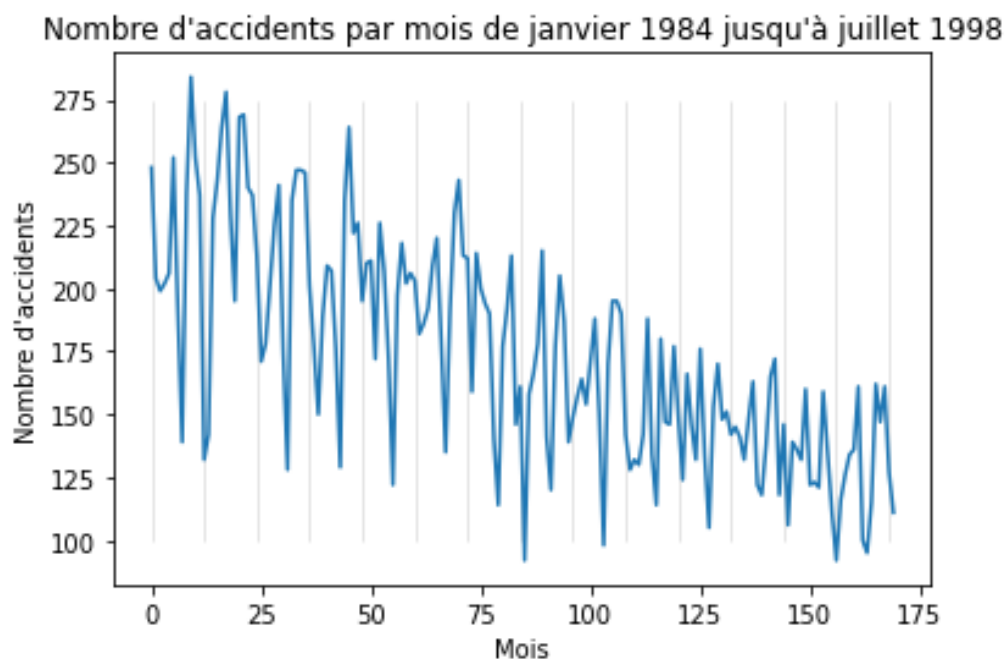
Dans la même démarche que la *Question 1*, nous voulions montrer une différence significative entre les moments où se produisent les accidents, mais cette fois ci à l'échelle d'une journée. Nous pensions dans un premier temps que la différence entre le nombre d'accidents les jours fériés et les jours normaux aurait naturellement effacé la courbe des jours fériés (auquel cas nous aurions fait une mise à l'échelle en prenant le minimum pour chaque type de jour et le maximum), mais ce n'était en réalité pas le cas.

Ces courbes nous montrent que les accidents les jours fériés commencent à se produire légèrement plus tard (9 heures contre 6 heures pour les jours « normaux »), confirmant la conclusion de la question 1 : les accidents semblent se produire plus souvent sur des trajets habituels.



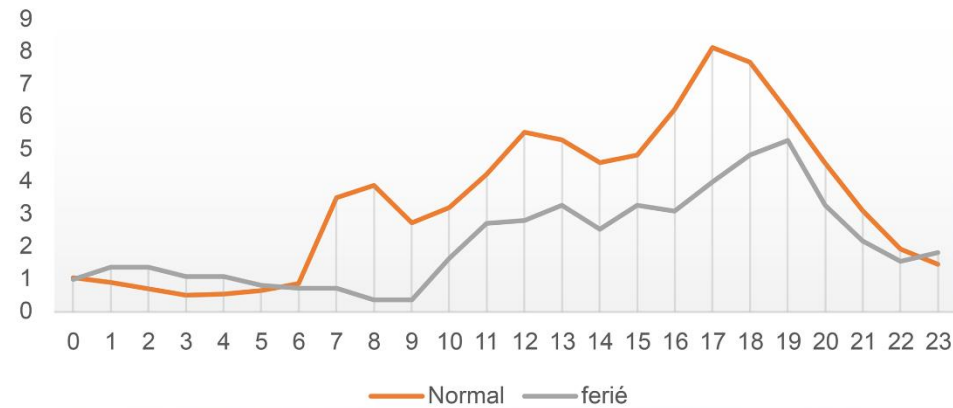
Autres graphiques :

Vous nous aviez demandé la réalisation d'un graphique montrant l'évolution des accidents sur les 14 années, mais nous en avons conclu, après sa réalisation, qu'il était hors-sujet car il ne suivait pas notre problématique. Nous pouvons nonobstant observer une baisse significative du nombre d'accidents, que l'on explique par une forte sensibilisation (1986 : retrait immédiat du permis en cas d'ivresse ; 1990 : contrôles d'alcoolémie et 50 km/h en ville ; 1995 : 0,5g/l d'alcool autorisé).



# ÉTUDE DES ACCIDENTS EN FONCTION DES JOURS PARTICULIERS ( VACANCES, PONTS, JOURS FÉRIÉS... )

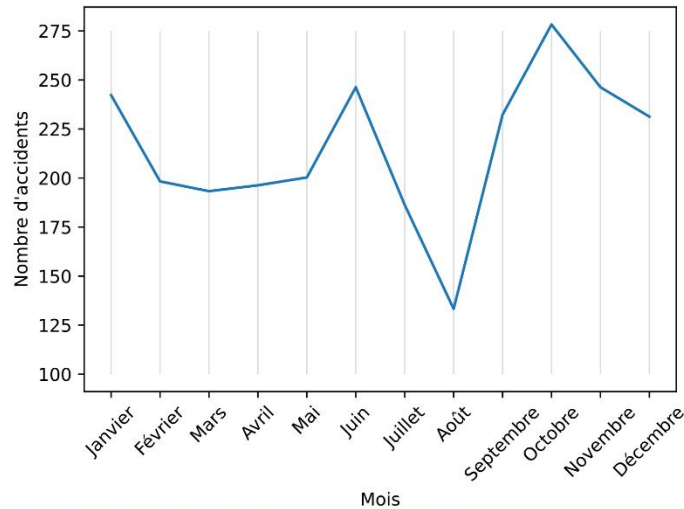
Nombre d'accidents moyen par heure de la journée



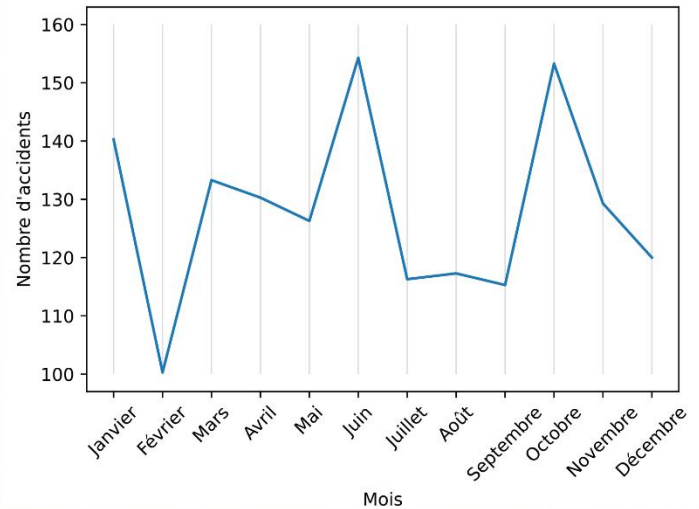
Moyenne des accidents par type de jour :

Tous	5,7113
« Normal	5,603
Férié	3,8987

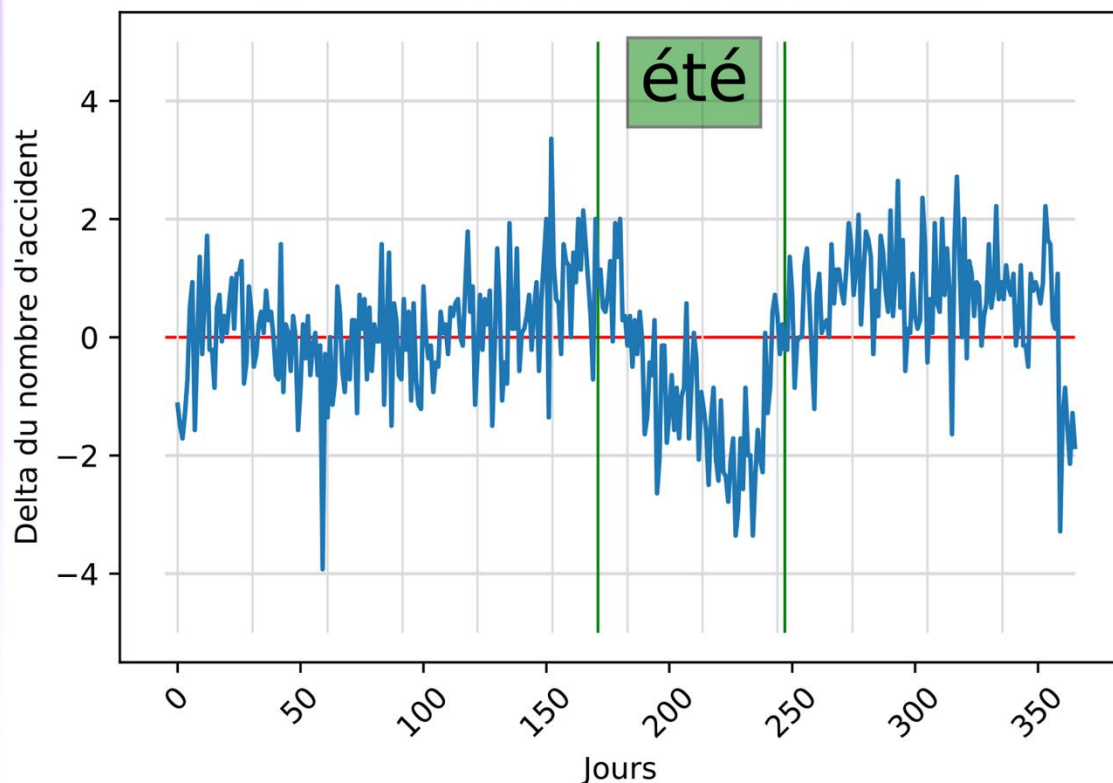
nombre d'accidents en fonction du mois dans l'année 1984



nombre d'accidents en fonction du mois dans l'année 1996



Delta du nombre d'accident par jour et la moyenne du nombre d'accident sur les 14 années



## Bilan :

Pour cette première partie de la S.A.É., nous pensons avoir acquis beaucoup de compétences concernant d'une part l'étude des données, puis de l'autre leur visualisation.

Pour ce qui est de l'étude des données, la création d'indicateurs est une étape importante pour la suite, car elle permet d'avoir des bases fiables pour réaliser des requêtes complexes. Ces indicateurs se doivent d'être pertinents. La conception de ces requêtes en fonction d'une thématique prédéfinie est aussi majeure car derrière en dépend la visualisation des données. Pour éviter les biais, mauvaises interprétations et mauvais affichage, il faut étudier au maximum le domaine qui s'offre à nous, ainsi le pipeline de l'exploitation de la base s'en trouvera grandement renforcé.

Pour ce qui est de la partie « Programmation », l'exploitation de cette base de données nous permet d'apprendre à manipuler les données sous MS Excel et sous Python, ainsi qu'à les représenter graphiquement selon des normes et avec le maximum de cohérence.

Enfin, pour exploiter une telle base de données, il faut trouver une organisation correcte afin de ne pas, par exemple, écraser malencontreusement les travaux déjà effectués.