

WebScrapping & Data Processing



Comparateur du prix de différents moyens de transport

Détermination du meilleur moment pour acheter ses billets, pour chaque type de transport



Contexte :

Le prix du moyen de transport dans le budget voyage est un élément important. On cherche à trouver le meilleur moyen de transport : train, avion ou bus, et ce au meilleur prix.

Étant donné que le coût élevé des billets de train et billets d'avion n'est pas quelque chose que nous pouvons ignorer lorsque nous planifions notre budget vacances, qui ne voudrait pas dépenser le minimum pour ses billets ? Seulement, il ne nous est pas toujours possible de faire une réservation lorsque les prix sont les plus bas, et surtout de savoir quand c'est le cas. De temps en temps, il y a de fortes baisses des prix des billets à des moments étranges.

Si on pouvait analyser et comprendre ces baisses de prix des billets, on aurait plus de chances de réserver les billets proches de notre date de voyage à des prix modestes. Et ainsi alléger le prix des transports dans notre budget voyage !

Idée du projet :

Récupérer les données de deux sites de réservation/achat de titres de transports pour voyager.

Comparer les prix de différents moyens de transports pour permettre à l'utilisateur de trouver la solution la moins chère parmi :


Bus, train, avion

Analyser l'évolution des prix pour ces moyens de transports afin de déterminer les moments optimaux lors desquels les prix sont les moins chers. Partager avec les utilisateurs les meilleurs moments pour acheter leurs billets au meilleur prix.

Premier site internet : Tictactrip

<https://www.tictactrip.eu/>



Transports Top trajets Contact Mes billets | 

Quel est votre trajet ?

Paris

Milan

 Lundi 16 janvier

 Dimanche 22 janvier

 1 adulte, 0 carte

Rechercher

Choisissez votre date de retour

< Janvier 2023 >

Lu	Ma	Me	Je	Ve	Sa	Di
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

À partir de

6h 8h 10h 12h 14h 16h 18h 20h 22h



Premier site internet : Tictactrip

<https://www.tictactrip.eu/>

The screenshot displays the Tictactrip website interface for a journey from Paris to Milan on January 16, 2023. The header shows the route and date, with a button to edit the search. Below this, a navigation bar lists dates from Saturday to Wednesday. A filter icon and a dropdown menu for sorting by time are visible. The main content area features a table of transport options: Train (starting at 59€), Bus (starting at 32€), Combined, and Carpooling. Two train options are listed with their departure times, direct status, and carbon footprint. A call-to-action button 'Aller' is present, along with a prompt to select a journey for more details.

Aller : Paris → Milano le 16/01/2023

13 trajets trouvés

Sam. 14 Janv. Dim. 15 Janv. Lun. 16 Janv. Mar. 17 Janv. Mer. 18 Janv.

Trier par horaires ▾

Train Dès 59 €	Bus Dès 32 €	Combiné -	Covoit' -
<p>06:39 Paris Gare de Lyon, Paris 7h10, Direct</p> <p>13:49 Milano Garibaldi, Milano</p> <p> Dès 59.00 €</p> <p> 1.10 kg CO₂</p>			
<p>12:43 Paris Gare de Lyon, Paris 7h06, Direct</p> <p>19:49 Milano Garibaldi, Milano</p> <p> Dès 64.00 €</p> <p> 1.10 kg CO₂</p>			

Aller

Sélectionnez un trajet pour accéder aux détails

Comm

Ce premier site, Tictatrip, propose des trajets en car et train dans toute l'Europe.

Par exemple pour ce rendre de Paris à Milan, nous avons la liste de tous les trajets disponibles (ci-contre affiché pour les trains). Le trajet le plus rentable en terme de durée/prix est mis en avant ainsi que le prix le plus avantageux.

Nous allons scraper les informations concernant les dates pour l'aller et le retour, les horaires, le temps de trajet, le type de transport, la compagnie assurant le voyage, le prix...

Second site internet : Opodo

<https://www.opodo.fr/>

opodo Vols Hôtels Vol + Hôtel Location de voiture Navette aéroport Autres [Gérer ma réservation](#) [Mon compte](#) [Mon abonnement Prime](#) [Info - Aide](#) [FR \(€\)](#)

Comparez et économisez

Économisez jusqu'à 40 %

☒ Vols ☐ Hôtels ☐ Vol + Hôtel ☐ Location de voiture

☒ Aller-retour ☐ Aller simple ☐ Multi-destinations ☒ Économique ☐ Vols directs uniquement

D'où partez-vous ? Où allez-vous ? Départ le Retour le 1 voyageur

[Rechercher Vol + Hôtel](#) [Rechercher les vols](#)

opodo prime Plus de 3 millions de membres Prime [Découvrez Prime](#)

3 MILLIONS de membres

[Gérer la réservation](#)
Modifiez ou annulez votre réservation, vérifiez le statut de votre vol et bien plus encore

[Vérifier le statut du remboursement](#)
Avez-vous une demande de remboursement en cours ? Vérifier le statut



Second site internet : Opodo

<https://www.opodo.fr/>

Ce second site, Opodo, propose des trajets en avion partout dans le monde.

Prenons le même exemple que pour le premier site, un voyage entre Paris et Milan. Nous avons une nouvelle fois une liste de trajets disponibles. Le trajet le plus rentable en terme de prix/durée est mis en avant ainsi que le moins cher et le plus rapide.

Nous allons scraper des informations similaires que celles scrapées sur le premier site.

The screenshot displays the Opodo website interface for a flight search from Paris to Milan. The top navigation bar includes the origin 'Paris', the destination 'Milan', the departure date 'lun. 16/01', the return date 'dim. 22/01', and a user icon with the number '1'. A 'Modifier' button is located on the right.

Below the navigation bar, there are three sorting options: 'Le meilleur' (1h 28 min - 63 €), 'Le moins cher' (9 h 13 min - 61 €), and 'Le plus rapide' (1 h 25 min - 127 €). A note states: 'Nos résultats sont classés par défaut selon 'le Meilleur'. Le classement des résultats peut être influencé par des critères contractuels ①'.

The main section is titled 'Meilleure option' and displays two flight options:

- ALLER - Vueling:** 18:40 from Orly (ORY) Paris to 20:05 at Malpensa (MXP) Milan. Duration: 1 h 25 min, direct. Price: 97,34 € (non réduit).
- RETOUR - Vueling:** 09:10 from Orlo Al Serio (BGY) Milan to 10:40 at Orly (ORY) Paris. Duration: 1 h 30 min, direct. Price: 62,92 € (Tarif Prime par passager).

A 'Sélectionner' button is present at the bottom right of the flight options section.

Below the flight options, there is a promotional banner for 'opodo prime' with the text: 'Profitez de centaines d'euros de réduction sur les vols, les hôtels et les voitures. Découvrir Prime →'.

At the bottom, another flight option is shown:

- ALLER - Easyjet:** 06:25 from Charles De Gaulle (CDG) Paris to 07:55 at Linate (LIN) Milan. Duration: 1 h 30 min, direct. Price: 26% de CO2 en moins (Plus que 9 billets!).

On the left side of the page, there are two bar charts showing CO2 emissions for flights from Paris to Milan. The top chart is labeled 'ions de CO2' and the bottom chart is labeled 'ions d'émissions de i2'. Both charts show a distribution of bars with a red dot at the end, indicating a specific value of 23:59.

Scraping : Train et Bus - Tictactrip

	Horaire_Depart	Horaire_Arrivee	ville_depart	ville_arrivee	Gare_Depart	Gare_Arrivee	Prix_Train	Duree_Train	Escale_Train
0	06:55	10:05	Paris	Stuttgart	Paris Est	Stuttgart	75.00 €	3h10	Direct
1	10:55	14:05	Paris	Stuttgart	Paris Est	Stuttgart	80.00 €	3h10	Direct
2	13:55	17:04	Paris	Stuttgart	Paris Est	Stuttgart	75.00 €	3h09	Direct
3	15:55	19:04	Paris	Stuttgart	Paris Est	Stuttgart	75.00 €	3h09	Direct
4	17:55	21:04	Paris	Stuttgart	Paris Est	Stuttgart	75.00 €	3h09	Direct

Nous avons utilisé Selenium pour scraper les informations sur Tictactrip.

Exemple ici : Dataset obtenu pour un aller pour un voyage de Paris à Stuttgart le 16 janvier.

Scraping : Avion - Opodo

	Ville_depart	Ville_arrivee	Date_depart	Date_retour	Nb_passagers	Nb_adultes	Nb_enfants	Nb_bebes	Aller_depart_aeroport	Aller_arrivee_aeroport	...
0	Paris	Stuttgart	16/01/2023	22/01/2023	5	2	2	1	Charles De Gaul... (CDG)	Echterdingen (STR)	...
1	Paris	Stuttgart	16/01/2023	22/01/2023	5	2	2	1	Charles De Gaul... (CDG)	Echterdingen (STR)	...
2	Paris	Stuttgart	16/01/2023	22/01/2023	5	2	2	1	Charles De Gaul... (CDG)	Echterdingen (STR)	...
3	Paris	Stuttgart	16/01/2023	22/01/2023	5	2	2	1	Charles De Gaul... (CDG)	Echterdingen (STR)	...
4	Paris	Stuttgart	16/01/2023	22/01/2023	5	2	2	1	Charles De Gaul... (CDG)	Echterdingen (STR)	...
5	Paris	Stuttgart	16/01/2023	22/01/2023	5	2	2	1	Charles De Gaul... (CDG)	Echterdingen (STR)	...
6	Paris	Stuttgart	16/01/2023	22/01/2023	5	2	2	1	Charles De Gaul... (CDG)	Echterdingen (STR)	...
7	Paris	Stuttgart	16/01/2023	22/01/2023	5	2	2	1	Charles De Gaul... (CDG)	Echterdingen (STR)	...
8	Paris	Stuttgart	16/01/2023	22/01/2023	5	2	2	1	Charles De Gaul... (CDG)	Echterdingen (STR)	...
9	Paris	Stuttgart	16/01/2023	22/01/2023	5	2	2	1	Charles De Gaul... (CDG)	Echterdingen (STR)	...
10	Paris	Stuttgart	16/01/2023	22/01/2023	5	2	2	1	Charles De Gaul... (CDG)	Echterdingen (STR)	...
11	Paris	Stuttgart	16/01/2023	22/01/2023	5	2	2	1	Charles De Gaul... (CDG)	Echterdingen (STR)	...

Nous avons également utilisé Selenium pour scraper les informations sur Opodo.

Retour_depart_heure	Retour_arrivee_heure	Retour_compagnie	Retour_duree_vol	Retour_escale	Retour_bagage	Prix
18:25	19:55	· Air France	1 h 30 min	direct	Bagage à main	157,95 €
18:25	19:55	· Air France	1 h 30 min	direct	Bagage à main	157,95 €
07:30	09:00	· Air France	1 h 30 min	direct	Bagage à main	157,95 €
11:30	12:50	· Air France	1 h 20 min	direct	Bagage à main	157,95 €
11:30	12:50	· Air France	1 h 20 min	direct	Bagage à main	157,95 €
07:30	09:00	· Air France	1 h 30 min	direct	Bagage à main	157,95 €
19:20	22:05	· Swiss International Air Lines	2 h 45 min	1 escale	Bagage à main	169,46 €
19:20	22:05	· Swiss International Air Lines	2 h 45 min	1 escale	Bagage à main	169,46 €
19:20	22:05	· Swiss International Air Lines	2 h 45 min	1 escale	Bagage à main	167,78 €
14:35	18:05	· Lufthansa	3 h 30 min	1 escale	Bagage à main	173,52 €
19:20	22:05	· Swiss International Air Lines	2 h 45 min	1 escale	Bagage à main	194,20 €
10:20	13:55	· Swiss International Air Lines	3 h 35 min	1 escale	Bagage à main	168,40 €

Scraping : Trajets étudiés

Jour du trajet étudié : 06/02/2023

Itinéraires :

- Paris-Lyon
- Paris-Stuttgart
- Paris-Budapest
- Bordeaux-Lyon
- Lyon-Marseille (seulement avion)

Jour des relevés :

4 au 18 février

Récupération des données pour chaque trajet :

- Chaque jour
- Chaque heure
- Sur deux semaines (dans l'idéal, sur plusieurs mois)

-> Comme nous devons lancer le script de scraping manuellement tous les jours, à chaque heure, il nous manque certains relevés, surtout entre 1h et 6h.

A1	Ville_depart,Ville_arrivee,Date_depart,Nb_passagers,Nb_adultes,Nb_enfants,Nb_bebes,Aller_depart_aeroport,Aller_arrivee_aeroport,Aller_depart_heure,Aller_arrivee_he														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Ville_depart,Ville_arrivee,Date_depart,Nb_passagers,Nb_adultes,Nb_enfants,Nb_bebes,Aller_depart_aeroport,Aller_arrivee_aeroport,Aller_depart_heure,Aller_arrivee_he,Aller_duree_vol,Aller														
2	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),08:55,10:10,10,0,0,Easyjet,1 h 15 min,direct,"90,99 à,-",13,1,2023,13														
3	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),20:45,21:55,1,0,0,Easyjet,1 h 10 min,direct,"76,55 à,-",13,1,2023,13														
4	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),06:05,07:15,1,0,0,Air France,1 h 10 min,direct,"136,96 à,-",13,1,2023,13														
5	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),07:00,08:15,1,0,0,Easyjet,1 h 15 min,direct,"85,34 à,-",13,1,2023,13														
6	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),18:35,19:50,1,0,0,Air France,1 h 15 min,direct,"163,99 à,-",13,1,2023,13														
7	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),13:35,18:05,1,0,0,Lufthansa,4 h 30 min,1 escale,"286,37 à,-",13,1,2023,13														
8	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),19:25,08:35,1,0,0,DiffA@rentes compagnies,13 h 10 min,1 escale,"67,89 à,-",13,1,2023,13														
9	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),22:55,13:15,1,0,0,Vueling,14 h 20 min,1 escale,"68,93 à,-",13,1,2023,13														
10	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),09:20,12:50,1,0,0,Air France,3 h 30 min,1 escale,"377,35 à,-",13,1,2023,13														
11	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),13:45,17:35,1,0,0,Air France,3 h 50 min,1 escale,"377,01 à,-",13,1,2023,13														
12	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),22:55,12:15,1,0,0,DiffA@rentes compagnies,13 h 20 min,2 escales,"130,39 à,-",13,1,2023,13														
13	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),11:05,13:15,1,0,0,DiffA@rentes compagnies,26 h 10 min,2 escales,"286,88 à,-",13,1,2023,13														
14	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),22:55,17:30,1,0,0,DiffA@rentes compagnies,18 h 35 min,2 escales,"146,33 à,-",13,1,2023,13														
15	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),22:55,10:25,1,0,0,DiffA@rentes compagnies,11 h 30 min,2 escales,"186,23 à,-",13,1,2023,13														
16	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Lyon Part Dieu ...,22:55,20:04,1,0,0,DiffA@rentes compagnies,21 h 09 min,2 escales,"257,25 à,-",13,1,2023,13														
17	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),13:20,13:15,1,0,0,DiffA@rentes compagnies,23 h 55 min,2 escales,"328,25 à,-",13,1,2023,13														
18	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),11:25,17:35,1,0,0,Air France,6 h 10 min,1 escale,"354,45 à,-",13,1,2023,13														
19	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),09:20,14:05,1,0,0,Air France,4 h 45 min,1 escale,"545,72 à,-",13,1,2023,13														
20	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),10:10,14:05,1,0,0,Air France,3 h 55 min,1 escale,"671,22 à,-",13,1,2023,13														
21	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),22:55,21:35,1,0,0,DiffA@rentes compagnies,22 h 40 min,2 escales,"244,92 à,-",13,1,2023,13														
22	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Lyon Part Dieu ...,22:55,12:03,1,0,0,DiffA@rentes compagnies,13 h 08 min,2 escales,"312,46 à,-",13,1,2023,13														
23	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),13:45,19:50,1,0,0,Air France,6 h 05 min,2 escales,"446,77 à,-",13,1,2023,13														
24	Bordeaux,Lyon,06/02/2023,1,Merignac (BOD),Saint Exupery (LYS),06:35,10:25,1,0,0,Air France,3 h 50 min,1 escale,"545,72 à,-",13,1,2023,13														
	Scrap data OPODO Bordeaux-Lyon														

On récupère pendant 2 semaines, toutes les heures, les résultats pour ces trajets pour le même jour de départ : par exemple pour un trajet Paris-Lyon le 6 février, on exécute notre script tous les jours à chaque heure pour récupérer les résultats et prix proposés.

Observations AVION

Evolution du prix moyen par heure pour un billet pour le 06/02/2023

Paris - Lyon



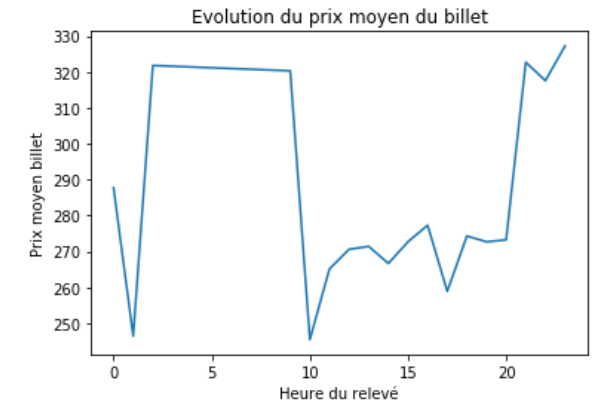
Paris - Stuttgart



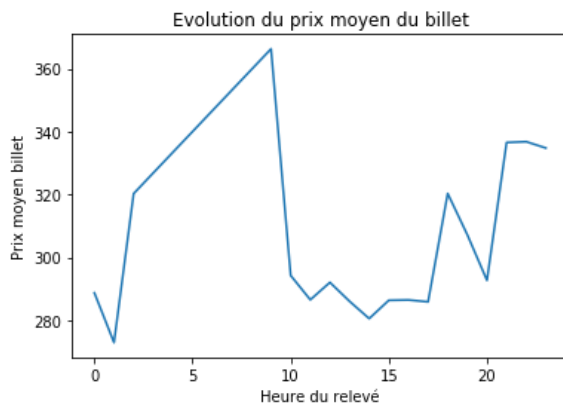
Paris - Budapest



Bordeaux - Lyon



Lyon - Marseille



On observe ici aussi, pour l'évolution du prix moyen par heure du billet, des comportements assez différents.

Pour les trajets Paris-Lyon et Paris-Stuttgart, il est plus intéressant de prendre ses billets avant 10h ou bien après 20h.

A l'inverse, pour les autres trajets, il est plus intéressant de prendre ses billets entre 10h et 20h.

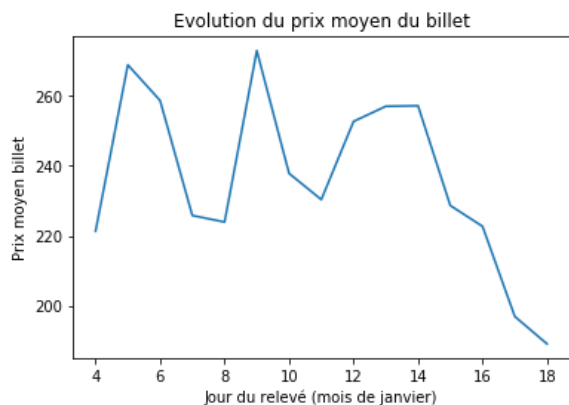
Les résultats qu'on obtient sont sûrement en partie biaisés par le fait que nous avons moins de données de relevé entre 1h et 6h.

Il est donc également COMPLEXE à ce stade de décrire un comportement généralisé pour l'évolution par heure du prix moyen du billet.

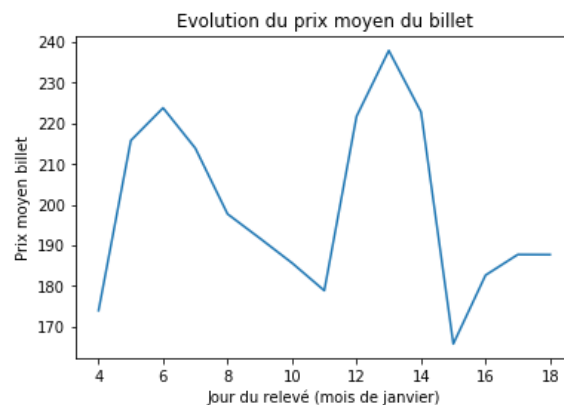
Observations AVION

Evolution du prix moyen par jour pour un billet pour le 06/02/2023

Paris - Lyon



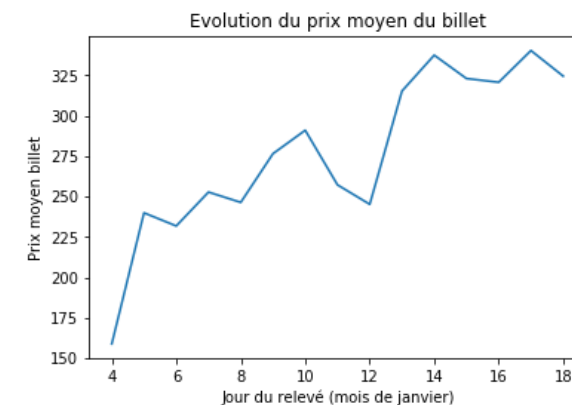
Paris - Stuttgart



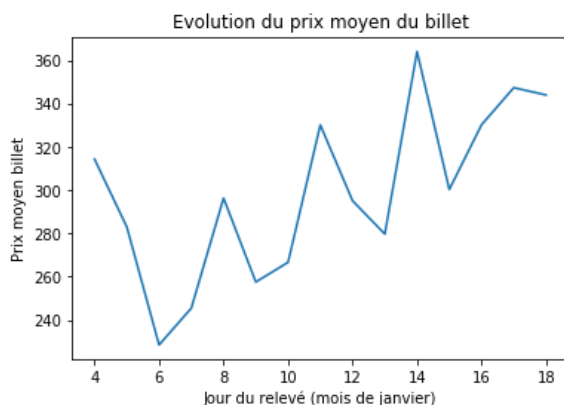
Paris - Budapest



Bordeaux - Lyon



Lyon - Marseille



On s'attendrait à ce que le prix moyen augmente progressivement lorsque l'on se rapproche de la date du vol. Cependant on observe un phénomène contraire pour les trajets Paris-Lyon et Paris-Stuttgart. Il serait donc intéressant de continuer à relever les prix pour les jours suivants la dernière date de relevé, à savoir le 18/01, pour voir si le prix continue à baisser ou bien remonte.

On observe pour ces trajets des pics lorsque la date du relevé est un mois avant la date du voyage : 06/01 pour le 06/02.

Pour le trajet Paris-Budapest, le prix moyen augmente puis semble stagner. Il rencontre seulement un faible pic pour le relevé du 14 janvier.

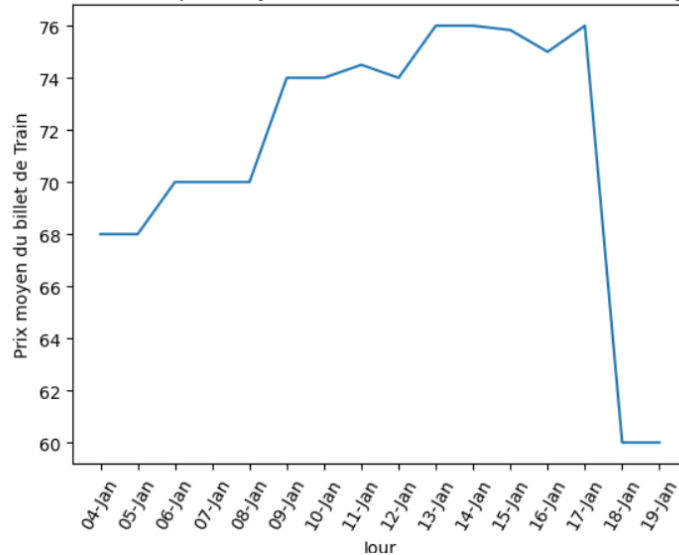
Pour les trajets Bordeaux-Lyon et Lyon-Marseille, on observe une augmentation progressive du prix moyen du billet à mesure que l'on se rapproche de la date du voyage. Pour le trajet Lyon-Marseille on observe cependant des pics d'augmentation du prix tous les 3 jours environ.

Il est donc COMPLEXE avec les données à notre disposition à ce stade de décrire un comportement généralisé pour l'évolution par jour du prix moyen du billet.

Observations TRAIN

Evolution du prix moyen par jour pour un billet pour le 06/02/2023

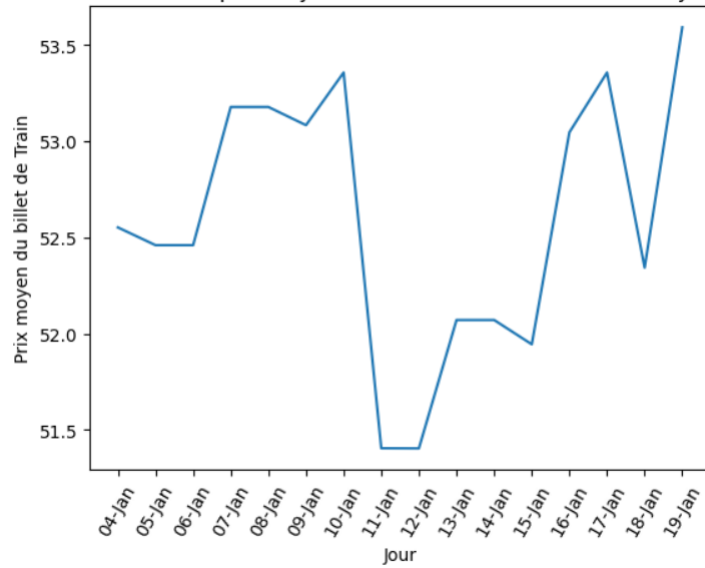
Evolution du prix moyen du billet de train entre Paris et Stuttgart



Parmi les 4 trajets que nous avons étudié des trajets en train n'étaient disponible que pour 2 d'entre eux, à savoir Paris-Lyon et Paris-Stuttgart. Nous pouvons imaginé que des destinations trop loin comme Budapest sont possibles en train mais avec un nombre de correspondance trop élevé pour que la trajet soit proposé.

Nous pouvons voir ci-contre les courbes d'évolution du prix de ces trajets en train pour un billet à la date du 06 février prochain.

Evolution du prix moyen du billet de train entre Paris et Lyon



Le prix moyen ne varie presque pas. Par exemple sur le trajet entre Paris et Lyon, en deux semaines, la différence entre le prix maximal et minimal n'est que de 2€.

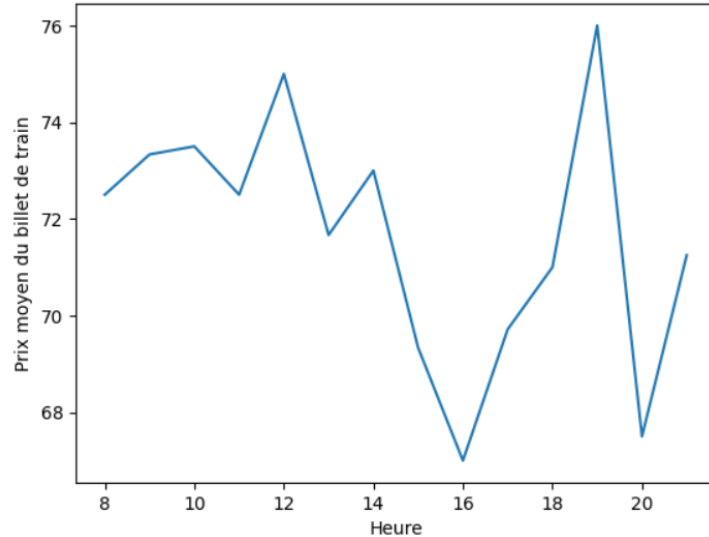
Pour le billet à destination de Stuttgart, nous pouvons constater une augmentation progressive du prix jusqu'à un pic de baisse.

De manière générale, nous ne pouvons pas définir de courbe ou de schéma visible. Tout comme pour l'avion, un comportement généralisé de ces prix pourraient être visible avec un dataset bien plus grand.

Observations TRAIN

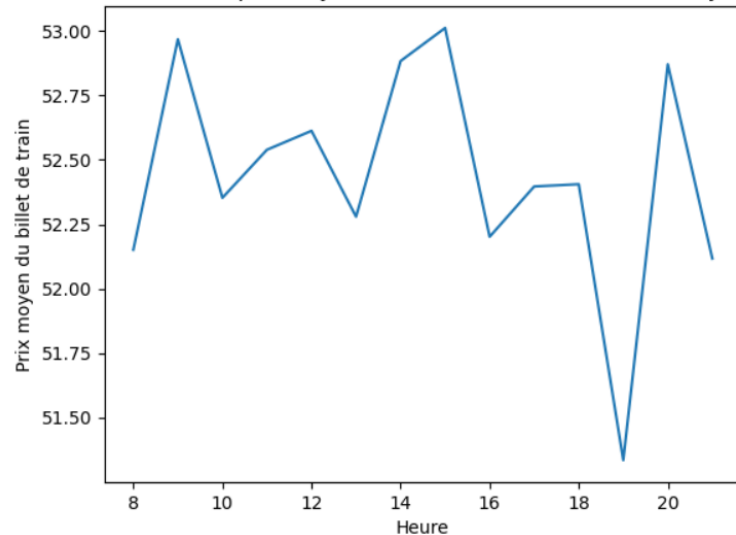
Evolution du prix moyen par heure pour un billet pour le 06/02/2023

Evolution du prix moyen du billet de train entre Paris et Stuttgart



Tout comme pour les prix jour par jour, nous ne voyons pas vraiment de schéma se répéter. Pour le trajet entre Paris et Stuttgart on peut voir deux piques vers le bas de la fourchette de prix à 16 et 20 heures.

Evolution du prix moyen du billet de train entre Paris et Lyon

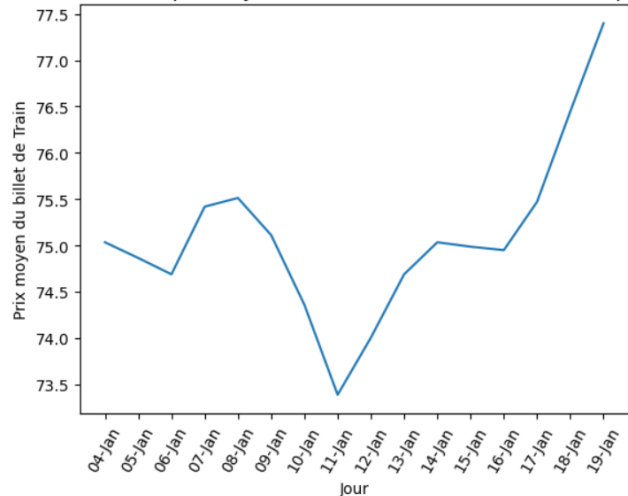


En revanche pour le trajet entre Paris et Lyon ce pique est à 19 heures, mais comme évoqué précédemment la différence de prix n'est que de 2€ donc pas assez significative.

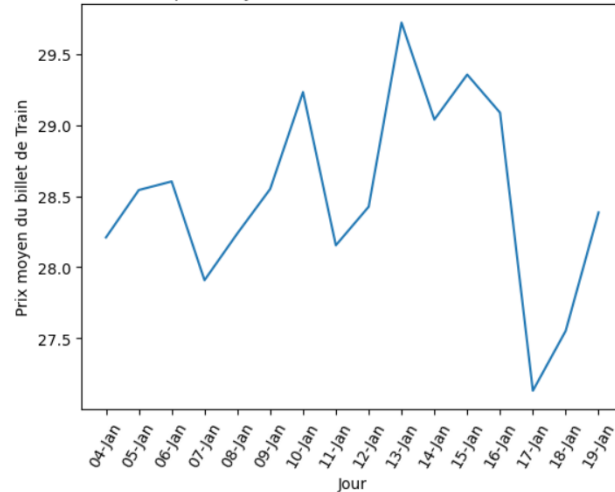
Observations BUS

Evolution du prix moyen par jour pour un billet pour le 06/02/2023

Evolution du prix moyen du billet de train entre Paris et Budapest



Evolution du prix moyen du billet de train entre Bordeaux et Lyon



Concernant le bus, moyen le moins cher de transport mais bien souvent le plus long, nous pouvons voir ci-contre l'évolution du prix des 4 voyages.

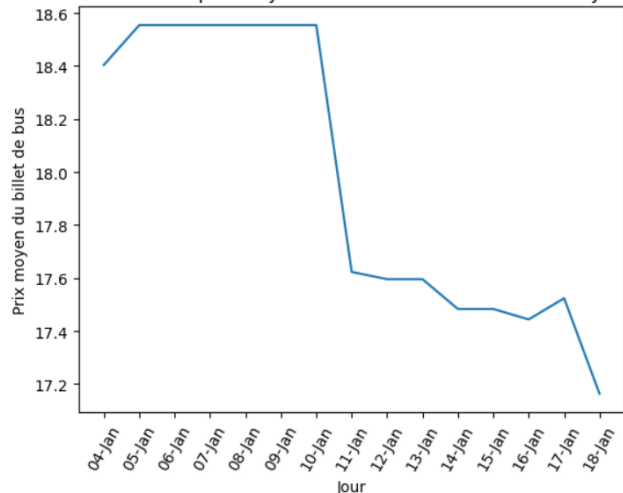
Tout comme pour les trains les schémas ci-contre sont un peu aléatoires et on ne peut pas identifier de schéma.

Pour le trajet entre Budapest et Paris ainsi que celui entre Paris et Stuttgart nous pouvons voir que le prix augmente selon que l'on se rapproche de la date du voyage.

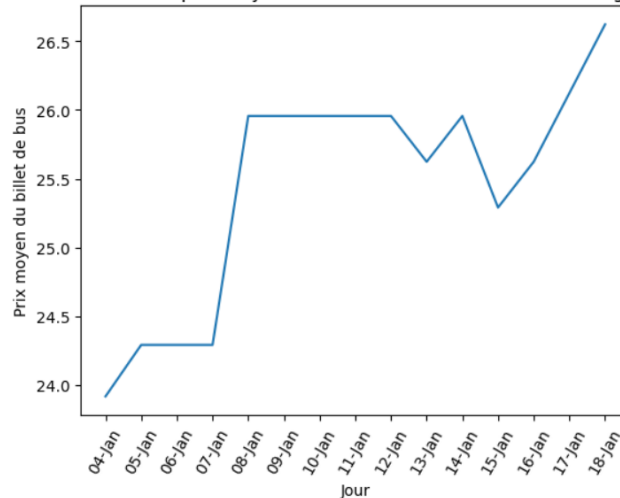
A l'inverse, le prix pour Lyon depuis paris baisse.

Cependant, toutes ces remarques ne sont pas valables puisque comme pour le train le prix ne change presque pas, seulement de quelques euros.

Evolution du prix moyen du billet de bus entre Paris et Lyon



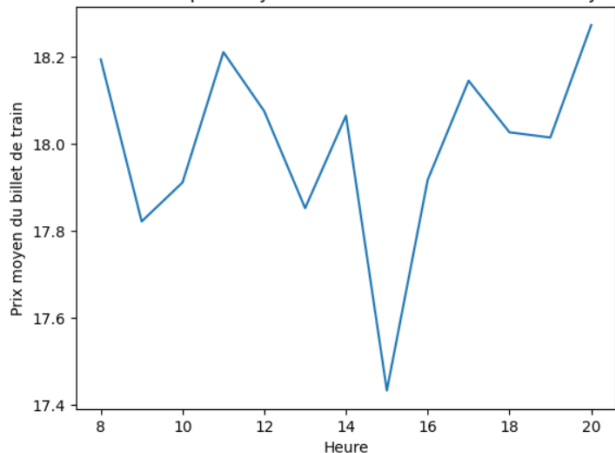
Evolution du prix moyen du billet de bus entre Paris et Stuttgart



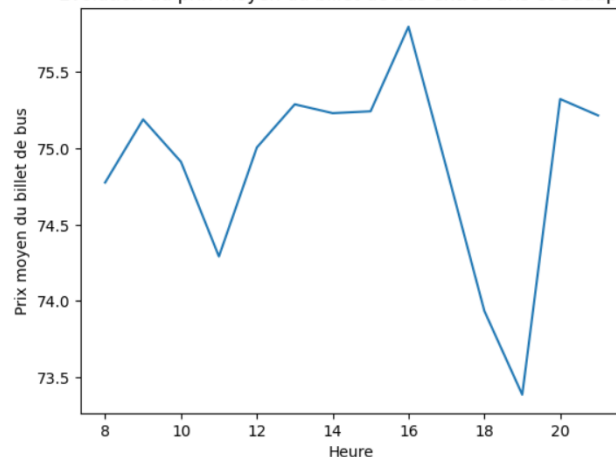
Observations BUS

Evolution du prix moyen par heure pour un billet pour le 06/02/2023

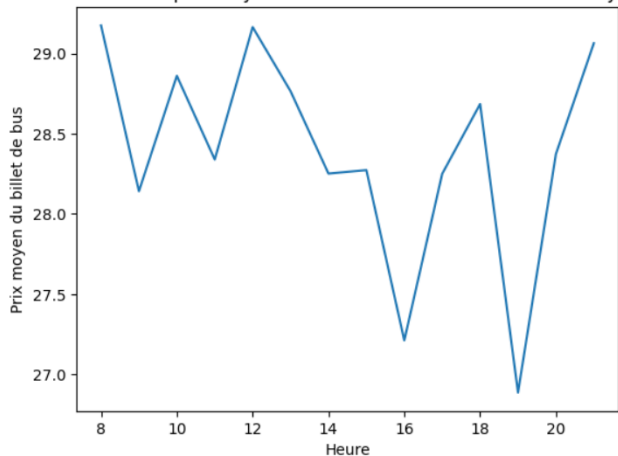
Evolution du prix moyen du billet de train entre Paris et Lyon



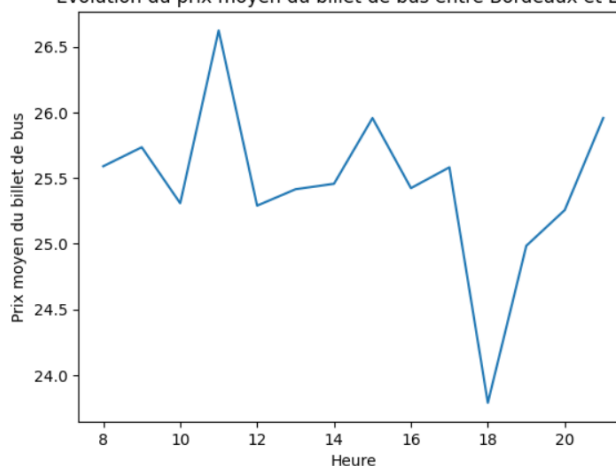
Evolution du prix moyen du billet de bus entre Paris et Budapest



Evolution du prix moyen du billet de bus entre Bordeaux et Lyon



Evolution du prix moyen du billet de bus entre Bordeaux et Lyon



On peut constater des piques de baisse de prix vers la fin de journée, mais encore une fois aucun schéma spécifique n'est visible. Surtout, la différence de prix est à nouveau trop faible pour être significative.

Conclusions - Avion

Nous n'avons pas eu des résultats très concluants pour déterminer le meilleur moment pour acheter ses billets d'avion. En effet, en fonction du trajet du vol, nous avons observé des comportements assez différents. Nous ne pouvons à ce stade pas généraliser ces résultats.

Il va falloir poursuivre le scraping en faisant des relevés à chaque heure, chaque jour, sans aucuns trous. Et répéter cela sur plusieurs semaines, mois. Mais aussi réaliser ces relevés pour davantage de dates, pas seulement une date pour pouvoir essayer de déterminer un comportement d'évolution général.

De cette manière, nous pourrons étudier les variables importantes qui ont une influence sur les prix des billets et en construire de nouvelles telles que « Week-end » (les billets les vendredi soir, samedi ou dimanche sont souvent plus chers), « Jour férié », « En période de vacances scolaires », « Saison », « Type de compagnie », etc.)

Nous pourrons ainsi construire un modèle pour déterminer le meilleur moment pour prendre son billet d'avion, pour l'avoir au meilleur prix. À ce stade, il ne nous est pas encore possible de construire un tel modèle.

Conclusions – Train et Bus

Les conclusions que l'on peut faire concernant nos observations sur l'évolution du prix sont similaires à celles pour l'avion.

Nous n'avons pas pu mettre en avant un comportement particulier, que ce soit par rapport au jour de la semaine, au weekend, à l'heure ou tout simplement à l'évolution du temps.

Tout comme pour l'avion il serait intéressant de pousser cette étude sur plusieurs mois pour voir si des schémas apparaissent vraiment ou si l'évolution des prix est vraiment aléatoire.

De plus, nous n'avons pas scrappé les données aussi souvent que nous l'aurions voulu pour éviter de se faire bannir des sites que nous avons utilisé. Une récolte des données faite de manière plus régulière pourrait également améliorer les résultats.

Conclusions Générales

Pour conclure sur ce projet, nous pouvons dire que le bus est le moyen le moins cher de voyager. Cependant la durée du trajet et les correspondances ne nous permettent pas de dire que c'est le moyen le plus rentable. Concernant l'avion, nous avons pu voir que le prix est imprévisible mais surtout que les différences de prix sont grandes et aléatoires. Pour finir, le train quant à lui ne change pas vraiment de prix et les temps de voyages sont bien plus courts que ceux en bus. De plus, le train reste le moyen de locomotion le plus écologique en termes d'émissions de CO2.

Techniquement parlant, nous ne sommes pas arrivées aux conclusions auxquelles nous nous attendions. Avec plus de temps et donc plus de matière, de relevés réalisés sur une beaucoup plus longue durée, nous aurions peut-être pu observer des schémas d'évolution des prix intéressants. Cela, dans le but de prévoir le moment opportun pour réserver son moyen de locomotion pour un voyage au prix le plus bas. Ou bien peut-être que l'évolution des prix pour les moyens de locomotion que sont le bus, le train et l'avion, est un phénomène trop aléatoire à prédire et modéliser.

Annexes

Opodo : <https://www.opodo.fr/>

TicTacTrip : <https://www.tictactrip.eu/>

Lien du github : https://github.com/marie-donier/Web_scraping_project