# Dossier élèves - Énoncé et contexte



La jeune pousse (*start-up*) rennaise BDDev - BigDataDeveloppements développe une nouvelle application mobile, disponible sur toutes les plateformes (notamment Google Store et Apple Store) nommée *cheapark* qui permettra à terme de connaître en temps réel les places de parking disponibles dans les grandes villes ou métropoles. BDDev s’est fait une spécialité de l’exploitation des données ouvertes des agglomérations, dont Rennes fut l’un des précurseurs en 2010.

Avec les données ouvertes des agglomérations, elle propose depuis 5 ans l’application *wifree* qui permettait au départ de connaître tous les réseaux wifi en libre accès dans la métropole rennaise, puis s’est étendue au fur et à mesure de l’accessibilité des données dans d’autres grandes villes comme Paris, Lyon ...

Le développement de l’application mobile *cheapark* repose sur l’exploitation de données de près de 40 agglomérations à commencer par les différents parkings implantés, leur localisation, leurs horaires d’ouverture, leur capacité, les types de véhicules admis... pour les parkings qui sont payants, les tarifs seront collectés ainsi que les modalités de paiement (carte, espèces, mais aussi via un badge de télépéage).

---

Vous travaillez au sein du groupe projet *cheapark* de la société BDDev qui a en charge le développement de la future application. Le travail de ce groupe projet consiste au recensement et l’analyse des données ouvertes des agglomérations / métropoles du territoire français, de la caractérisation des données nécessaires à l’application en développement. Votre travail consiste à collaborer avec l’équipe sur les premières étapes du projet.

## Première partie – Découverte des données ouvertes de diverses sources

Les données ouvertes sont accessibles pour un grand nombre de domaines, notamment les données sur les territoires, l’éducation, la population, les infrastructures. Chaque domaine propose des jeux de données qui peuvent être visualisées directement ou téléchargées sous divers formats d’échanges.

Le site gouvernemental <https://www.data.gouv.fr> centralise les principales données ouvertes, en donnant des exemples de réutilisation et valorisation comme par exemple le site mobe.fr qui, à partir de données ouvertes, permet de découvrir toutes les ventes immobilières d’un quartier avec une simple recherche.

Une image contenant texte, carte

Description générée automatiquement  
*Site mobe.fr*

Par exemple, dans le domaine « *Agriculture et Alimentation* », sur le thème de l’alimentation, il existe des données ouvertes permettant de connaître tous les ingrédients, informations nutritionnelles et données sur les produits alimentaires :

Une image contenant capture d’écran

Description générée automatiquement  
*Site www.data.gouv.fr*

### Travail à faire :

À partir de l’exemple précédent :

1. Donner au moins deux exemples d’utilisation possible des données sur les produits alimentaires (par exemple, information aux personnes présentant des allergies...).

2. Indiquer, pour chaque exemple d’utilisation possible trouvé, sous quelle(s) forme(s) (site web, application mobile, dépliants, publicités, annonces sur des médias...) ces données peuvent être transmises à un public de personnes ciblées (par exemple, une affiche dans une salle d’attente médicale).  
*Il est possible de traiter cette question sous la forme d’un tableau.*

Le Ministère de l’éducation nationale et de la jeunesse produit également des données ouvertes disponibles sur le site [https://data.education.gouv.fr](https://data.education.gouv.fr/)

### Travail à faire :

À partir du site [https://data.education.gouv.fr](https://data.education.gouv.fr/) :

1. Dans les thèmes « élèves », « examens » et « personnels » trouver pour chacun un exemple de jeu de données et repérer la dernière actualisation de ces données.

Les principales métropoles françaises publient leurs données ouvertes, qui sont recensées dans le tableau qui suit :

|  |  |
| --- | --- |
| Paris | <https://opendata.paris.fr> |
| Lyon | [https://data.beta.grandlyon.com](https://data.beta.grandlyon.com/) |
| Montpellier | https://data.montpellier3m.fr/ |
| Toulouse | https://data.toulouse-metropole.fr/ |
| Bordeaux | https://data.bordeaux-metropole.fr |
| Aix-Marseille-Provence | <https://www.agglo-paysdaix.fr/opendata-telecharger-les-donnees.html> |
| Rennes | [https://data.rennesmetropole.fr](https://data.rennesmetropole.fr/) |
| ... | ... |

Pour la plupart des métropoles citées ci-dessus ces données sont libres d’accès et peuvent être extraites directement du site, pour d’autres comme Aix-Marseille il faut en faire la demande.

Une image contenant capture d’écran

Description générée automatiquement

*Site de la ville de Paris* [*https://opendata.paris.fr*](https://opendata.paris.fr)

Une image contenant capture d’écran

Description générée automatiquement

*Site de la métropole de Lyon* [*https://data.beta.grandlyon.com*](https://data.beta.grandlyon.com/)

Une image contenant capture d’écran

Description générée automatiquement

*Site de la métropole de Rennes* [*https://data.rennesmetropole.fr*](https://data.rennesmetropole.fr/)

### Travail à faire :

À partir des données d’une métropole de votre choix :

1. Recenser les différents domaines couverts par les données.

2. Indiquer quels mots clés caractérisent les données d’un domaine de votre choix.

Un grand nombre de villes, départements ou régions proposent également des données ouvertes :

|  |  |
| --- | --- |
| Ville de Suresnes | https://www.suresnes.fr/ma-ville/services-municipaux/open-data/ |
| Ville de Combs-la-Ville | https://data.combs-la-ville.fr |
| Ville de Savigny-le-Temple | http://www.savigny-le-temple.fr/content/portail-open-data |
| Département de Loire Atlantique | https://data.loire-atlantique.fr |
| Département de Seine Saint Denis | https://data.seine-saint-denis.fr |
| Département des Hauts de seine | https://opendata.hauts-de-seine.fr |
| Région Île-de-France | https://data.iledefrance.fr |
| ... | ... |

### Travail à faire :

À partir de votre environnement (ville, département, région) :

1. Trouver des données ouvertes des villes, départements les plus proches.

À partir des données de trois entités différentes (ville, métropole, département ou région) de votre choix :

2. Comparer les thèmes couverts et établir des correspondances.

3. Présenter dans un tableau les différents thèmes, en indiquant les entités qui proposent des jeux de données sur ceux-ci.

Toutes les données présentées ici, et utiles pour le projet de développement de la future application, sont des données ouvertes provenant du secteur public.

### Travail à faire :

1. Rechercher des données ouvertes fournies par le secteur privé, en indiquant leurs sources, le ou les domaines couverts, leur disponibilité.

2. Expliquer les raisons pour lesquelles les données du secteur public sont plus faciles à trouver que celles du secteur privé ?

Une image contenant hache

Description générée automatiquementVous avez maintenant une bonne vision des mégadonnées et des données ouvertes. Deux nouveaux membres ont intégré l’équipe du projet *cheapark* et vous êtes chargé d’une présentation sur les mégadonnées et les données ouvertes. Pour diffuser les informations essentielles notamment celles de veille informationnelle faite au sein de l’entreprise, un compte Twitter est utilisé ainsi qu’un hashtag #veille. Des informations doivent être transmises sur les thèmes des mégadonnées #BigData et des données ouvertes #OpenData.

### Travail à faire :

1. Préparer deux tweets pour sensibiliser l’équipe à l’usage mégadonnées et des données ouvertes en indiquant un lien vers un site de confiance sur ces sujets.

## Seconde partie – Analyse de la structure des données de stationnement

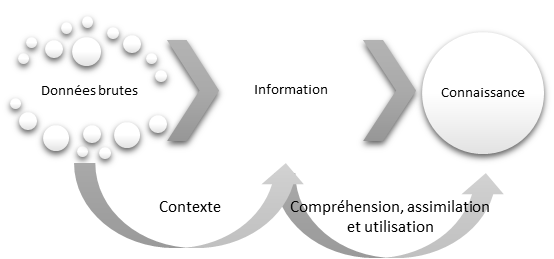
Vous vous intéressez à présent aux données relatives au projet de développement *cheapark* en cours. En effet, il faudra exploiter les données de stationnement des villes. Celles-ci sont disponibles selon diverses structures comme par exemple sur le site de la ville de Paris.

Une image contenant capture d’écran

Description générée automatiquement  
*Structure des données, représentations (tableau, graphique), valorisation  
Site de la ville de Paris* [*https://opendata.paris.fr*](https://opendata.paris.fr)

C’est en exploitant les données ouvertes que BDDev les interprète et les transforme en information puis en connaissance.

On peut représenter la transformation des données en information puis en connaissance, de la façon suivante :



*Schéma Données – information - connaissance*

Les données représentent un enjeu majeur pour les agglomérations notamment celles relatives au stationnement comme le précise les documents ci-dessous.

**Document 1 : La data au cœur du parking intelligent**

Une image contenant bâtiment, extérieur, route, rue

Description générée automatiquement

*Dans des villes de plus en plus embouteillées, le stationnement est un maillon clef de la mobilité. Au-delà des données statiques, de plus en plus de données dynamiques sont disponibles en open data, notamment grâce au développement des capteurs. Des données qui intéressent les acteurs publics comme privés, avec pour perspective de proposer l’offre de stationnement la plus complète possible aux usagers.*

En France, on estime que 15 à 30% de la circulation en ville est due à la recherche de places et 70 millions d’heures seraient perdues de ce fait… Si le sujet touche les grandes villes, il est également critique dans les villes moyennes où le véhicule individuel reste le principal moyen de locomotion. Mieux exploiter les places de parking vides, réduire le trafic lié aux véhicules en recherche de places, « désiloter » un monde marqué par une grande variété d’acteurs, du particulier aux collectivités en passant par les grands gestionnaires de parcs… le stationnement est au cœur de la « smart city » et d’un foisonnement d’initiatives publiques et privés. Leur point commun ? Utiliser la data, et de plus en plus d’open data, pour optimiser l’offre et enrichir le service aux automobilistes, de la recherche au paiement.

On notera aussi que la “datification” du stationnement est aussi poussée par la réforme du stationnement (loi Maptam) qui incite à une dématérialisation des processus de bout en bout, de la recherche de place à la gestion des FPS[[1]](#footnote-1).

**Caractéristiques des parkings et zonages**

Côté collectivités, le premier chantier a consisté à libérer les données « statiques » du stationnement en voirie et en ouvrage. Une requête sur data.gouv.fr montre plus de 300 résultats sur le mot-clef « stationnement » avec des villes de taille très diverses. Localisation des parkings, zonages tarifaires, emplacement des places PMR et des horodateurs figurent parmi les données les plus couramment publiées. Des données très appréciées des calculateurs d’itinéraires et plateformes cartographiques (Mappy, TomTom, Here, OpenStreetMap) des services spécialisés comme Jaccede.com ou l’appli “stationnement handicapé”, une application comme Parking gratuit ou un acteur international comme Parkopedia. Dans le cadre de la loi réforme du stationnement, les collectivités partagent aussi les données liées aux zonages / tarifications avec par exemple la plateforme créée par Versailles, qui cible dans un premier temps les prestataires européens (ou chinois) proposant de réserver des places dans cette ville hyper touristique.

**Disponibilité des places en temps réel**

Plus complexes à délivrer – puisqu’elles nécessitent des systèmes de comptages/ liaisons radios – les données de disponibilité des parkings en open data sont moins fréquentes. On les trouve par exemple sur le site de la Roche-sur-Yon et sur les portails data de métropoles comme Nantes, Lille, Bordeaux, Lyon ou le site de la Saemes. Ces données constituent une aide à la décision pour l’automobiliste (probabilité de trouver un stationnement) directement accessible sur leur mobile via des applications comme Nantes dans ma poche ou OPnGO. Faciles à collecter pour le stationnement en ouvrage, l’ambition de nombreuses collectivités est de proposer une offre de « smart parking » sur toutes les places de stationnement sur voirie. Des capteurs au sol ont été testés par des villes comme Nice, Toulouse ou encore Sète mais avec un succès mitigé. Ces capteurs prospèrent surtout dans les parkings privés, grandes surfaces ou aéroports (avec parfois des données en open data). Les villes semblent aujourd’hui leur préférer l’analyse d’images vidéo, les caméras pouvant aussi être utilisées pour le contrôle du stationnement payant. Soissons déploie ainsi la solution de ParkingMap qui analyse l’occupation de la voirie grâce à des boitiers vidéo doté d’un algorithme d’analyse d’images. Avec un usage double : la réalisation d’analyses quantitatives (durée, profil des véhicules stationnés) et l’occupation des places en temps réel via une application gratuite.

**Proposer une vision complète de l’offre**

Le stationnement ne se limite pas aux places en ouvrage ou sur la voie publique. Bailleurs, hôtels, particuliers… partagent de plus en plus leurs places de parking inoccupées. Avec à la clef un foisonnement de startups : Prendsmaplace.fr, Monsieurparking et Parkadom (entre particuliers), Bepark.eu (entreprises), Yespark.fr et Zenpark.com (qui ciblent les bailleurs, les gares, les administrations…). Ces sites, qui se rémunèrent par des commissions, misent sur une visibilité nationale, des tarifs dégressifs ou encore la commercialisation de dispositifs de contrôle de l’usage effectif du parking. Ces startups n’ont cependant qu’une offre limitée à un segment de marché, aussi pour proposer un service plus complet, elles utilisent de l’open data. Ces startups ouvrent parfois leurs données à des tiers : Mappy utilise ainsi les API de plusieurs entreprises (OPnGO, Saemes…). Ces données intéressent aussi les collectivités, comme à Saclay où le projet Moov’Hub mené avec plusieurs startups, vise à fournir à l’usager une information exhaustive sur l’offre locale de stationnement, publique ou privée.

*Rédacteur : Olivier Devillers, pour OpenDataFrance*<http://www.opendatafrance.net/2019/04/24/la-data-au-coeur-du-parking-intelligent/>

**Document 2 : Entretien avec Valérie Eynaud, chef de projets digital du Saemes**

Une image contenant personne, mur, intérieur, habits

Description générée automatiquement

**Quelles données proposez-vous en open data ?**

Saemes, délégataire de service public qui gère 90 parkings à Paris en Ile-de-France, est le premier acteur du stationnement à s’être lancé dans l’open data dès mars 2016. Notre plateforme propose 14 jeux de données publics : caractéristiques des parcs, horaires d’ouverture, tarifs, points de recharge électrique, accessibilité personnes à mobilité réduite… Avec également des données temps réel sur la disponibilité des places de parking ou celle des casiers des consignes à casques destinées aux deux roues. Toutes les données sont accessibles en téléchargement, sous fichier excel / csv / pdf, sous forme de carte ou d’API via opendata.saemes.fr. L’API génère environ 400 000 appels par mois.

**Comment sont réutilisées vos données ?**

C’est toujours la grande question... hélas nous avons une connaissance fort imprécise des réutilisations, sauf quand les réutilisateurs nous appellent. Kazakar, accessible.net, TomTom, ParkingMap et la ville de Gentilly (Saemes exploite les horodateurs de la ville) en font partie... sans oublier nos services puisque la plateforme est biface : elle a un volet grand public mais aussi une partie privée qui permet de partager des données utiles à différents services internes. Pour favoriser les échanges avec les réutilisateurs, nous avons créé une limite de 50 000 appels de l’API, seuil au-delà duquel on leur demande de prendre contact avec nous. Il ne s’agit pas de brider le dispositif mais d’animer l’écosystème et de mieux connaître les usages de nos données.

**Quels points vous paraissent devoir être améliorés ?**

Nous sommes en discussion avec la Mairie de Paris qui publie sur son site les données – statiques – de l’ensemble de l’offre de stationnement sur la capitale. Aujourd’hui cette source semble privilégiée sur la nôtre par beaucoup d’acteurs car plus complète en termes d’offre de stationnement bien que les données soient statiques. C’est assez logique. Avec un bémol cependant : les données ne sont pas à jour car fondées sur une ressaisie de documents fournies par Saemes dans le cadre du contrôle de la Délégation de Service Public ! Mais au-delà de ce problème, en voie de résolution, il faut que tous les acteurs du stationnement passent à l’open data. En effet, faute d’exhaustivité, nous n’intéressons pas des City Mapper et les autres poids lourds de la mobilité connectée. Le fait qu’Indigo – acteur mondial du stationnement – travaille à l’ouverture de ses données est à cet égard une bonne nouvelle. Cela va cependant enclencher d’autres sujets comme celui de la normalisation des données.

*Propos recueillis par Olivier Devillers, pour OpenDataFrance*<http://www.opendatafrance.net/2019/05/10/entretien-avec-valerie-eynaud-chef-de-projets-digital-du-saemes/>

### Travail à faire :

En exploitant les deux documents précédents et à l’aide de vos connaissances personnelles, répondre aux questions qui suivent :

1. Donner une information produite à partir d’un jeu de données de stationnement.

2. Donner un exemple de connaissance issue de cette information produite.

|  |
| --- |
|  |

### Travail à faire :

À partir du site web des données ouvertes de la ville de Paris :

1. Rechercher les jeux de données proposées relatifs au stationnement et observer les différents formats des fichiers contenant les jeux de données...

a) Recenser les différents formats des fichiers proposés en téléchargement.

b) Rechercher sur Internet les particularités de ces formats, qu’ont-ils en commun ?

2. Choisir le jeu de données « Horodateurs - mobiliers » afin de traiter les points qui suivent :

a) Lister quelques attributs intéressants disponibles dans ce jeu (un attribut est un nom décrivant une information de la base, c’est en quelque sorte le titre des colonnes contenant les données – exemple : attribut « prénom », données : « Jean », « Martine » ...).

b) Repérer le nombre d’enregistrements (occurrences) contenus dans ce jeu de données.

c) Dans le menu « tableau » et à l’aide de filtres, déterminer le nombre d’horodateurs solaires situés dans le premier arrondissement.

3. Indiquer quel est l’intérêt de décomposer la description d’un horodateur en plusieurs attributs dits « atomiques » (ne contenant qu’une seule valeur) ?

4. Préciser si tous les attributs de ce jeu de données sont atomiques ?

## Troisième partie – Exploitation des données de stationnement

Un des objectifs du projet *cheapark*, pour se faire connaître progressivement, est de publier un site applicatif zerocreneau.fr, recensant les stationnements de Paris où le stationnement en créneau peut être évité. Cela permettra de tester une des fonctionnalités de la future application et de recueillir les avis des utilisateurs en prévision de la publication des applications sur les différents stores.

Le site web est quasiment achevé, il faut insérer une carte et quelques informations utiles aux usagers.

Une image contenant capture d’écran

Description générée automatiquement  
*Interface du site web zerocreneau.fr en construction*

### Exploitation via le site opendata.paris.fr

##### Intégration et publication de la carte interactive

### Travail à Faire :

À partir du jeu de données « Stationnement sur voie publique – emplacements » de la ville de Paris ([https://opendata.paris.fr](https://opendata.paris.fr/)) :

1. Appliquer des filtres sur le « type de stationnement » de façon à avoir des données exceptées celles des créneaux (Longitudinal).

2. Insérer la carte dans le projet de site en suivant les étapes :

a) Ouvrir le dossier contenant le site web zerocreneau.fr et double-cliquer sur le fichier « index.html » pour le visualiser dans le navigateur.

b) Repérer visuellement l’emplacement de la future carte interactive sur le site.

c) Ouvrir, en parallèle, le même fichier « index.html » avec un éditeur de texte (bloc-notes, notepad ++ …) :

d) Repérer le numéro de ligne sur laquelle sera inséré le code html (langage utilisé pour structurer une page web) de la future carte interactive du site.

e) Repérer et copier le code d’intégration html (dans l’onglet « Carte »).

f) Coller le code dans le fichier « index.html »

g) Modifier le code collé afin que la carte ait une largeur de 100% et une hauteur de 600 pixels et sauvegarder.

h) Visualiser la carte intégrée dans le navigateur.

##### Création et publication d’informations diverses

### Travail à faire :

À partir du jeu de données « Stationnement sur voie publique – emplacements » de la ville de Paris ([https://opendata.paris.fr](https://opendata.paris.fr/)), dans l’onglet « Analyse » :

1. Générer un graphique permettant de dénombrer (compter) le nombre de stationnements par arrondissement.

2. Ventiler les séries selon le type de stationnement (bataille et épi).

3. Intégrer ce graphique sur le site web à l’emplacement adéquat (même mode opératoire que pour la carte).

### Exploitation via un tableur

Les données peuvent également être traitées et représentées grâce à un tableur en utilisant le format d’échange CSV.

### Travail à faire :

1. Exporter « seulement les *x* enregistrements sélectionnés » précédemment au format CSV et importer les données dans un tableur.

2. Générer un « Graphique croisé dynamique » permettant de dénombrer (compter) le nombre de types de stationnements par arrondissement (sélection de la cellule A1 -> Insertion -> Graphique croisé dynamique).

3. Indiquer quel est le principal intérêt d’utiliser un tableur ?

## Quatrième partie - Aller plus loin...

Le projet *cheapark* n’est qu’un exemple de l’exploitation des données ouvertes par les organisations. La mise à disposition des données par les organisations territoriales, les principaux services publics s’amplifie et peut permettre l’information du public tout comme créer des opportunités pour certaines entreprises.

### Travail à faire :

1. En une à deux pages, à partir de vos connaissances et en vous appuyant sur diverses situations de gestion dont celle présentée dans les trois premières parties, répondre de façon cohérente et argumentée à la question suivante :

**Les données ouvertes peuvent-elle être une ressource avantageuse pour les organisations ?**

En vous inspirant des travaux réalisés dans le cadre de cette ressource et de cet article https://www.nouvelobs.com/nos-regions-connectees/20190318.OBS1990/le-big-data-en-cinq-chiffres.html :

2. Réaliser une infographie (par exemple avec https://piktochart.com/ ou https://www.canva.com/) présentant les données ouvertes relative au domaine du stationnement.

1. Le forfait de post-stationnement (FPS) est une redevance forfaitaire de stationnement en France. [↑](#footnote-ref-1)