Project Swarm

# Outline

1. Version 0.1
2. 10/04/2019
3. Team: **Ilia Kassianenko, Luigi Clerici**

**Name of people in the team:** Ilia Kassianenko, Luigi Clerici

**Github repository:** <https://github.com/iliakass/CEB1260_repo/tree/master/Project>

**Dataset description:**

Parasite reports, based on extermination, within the nineteen borough of the City of Montreal are logged within the dataset. The City of Montreal anonymized the data by using intersections rather than physical addresses for origin of the report.

The data dictionary includes a unique ID of the extermination report, the date of the report, the date of the pre-extermination inspection, the number of extermination per report with a maximum of four, the date of the first extermination associated to the report, and the date of the last extermination associated to the report. The longitude and latitude coordinates are available for the referenced intersection per report. The reference number for the borough, its name, and the name of the neighbourhood are also available.

**Link for Dataset of interest:** See also references

<http://donnees.ville.montreal.qc.ca/dataset/declarations-exterminations-punaises-de-lit/resource/6173de60-c2da-4d63-bc75-0607cb8dcb74>

**Potential problems with dataset:**

* The dataset source files need to be extract in UTF-8 format in order to have French accents (ex. Saint-Léonard instead of Saint-LÃ©onard)

The chosen dataset has a low degree of reliability for its contents, which are provided by third parties such a parasite management firms. (Ville de Montréal 2019) These contents are not validated by the city of Montreal. (Ville de Montréal 2019)

* Pre-processing will be needed to determine whether the date format was entered correctly. Specifically, the following cases may occur.
  + First, it is possible that the month and day were recorded in reverse order by the third party. Although, this possibility cannot be verified expect by identifying a pattern and reviewing outliers for months and days.
  + Second, it possible latitude and longitude coordinates for street intersection were subject to interpretation, as the third party anonymized the actual locations through the use of street intersections. Specifically, each subset sources forming the third party may using different references for determining whether an event near one intersection rather than the opposite intersection.

# Bibliography

Ville de Montréal. *Déclarations des exterminations de punaises de lit par des gestionnaires de parasites.* February 08, 2019. http://donnees.ville.montreal.qc.ca/dataset/declarations-exterminations-punaises-de-lit (accessed April 09, 2019).

—. *Déclarations exterminations punaises de lit.* January 1, 2019. http://donnees.ville.montreal.qc.ca/dataset/declarations-exterminations-punaises-de-lit/resource/6173de60-c2da-4d63-bc75-0607cb8dcb74 (accessed April 9, 2019).

—. *Nombre de logements traités par quartier de référence.* January 1, 2019. http://donnees.ville.montreal.qc.ca/dataset/declarations-exterminations-punaises-de-lit/resource/81a05c27-2563-4cf5-9c40-ffcc053e05de (accessed April 9, 2019).