

# Labo 7: Prise en main avec BlueSky

Visseho Adjiwanou, PhD.

23 February 2022

## Introduction

Utiliser les données de l'enquête canadienne sur le revenu de 2015 pour répondre aux questions suivantes. On se limitera à une base de données réduites de cette base. Les variables retenues dans cette base sont:

- PERSONID: Person identifier (identifiant)
- AGEGP: Person's age group as of December 31 of reference year (groupe d'âge)
- SEX: Sex (Sexe du répondant)
- MARST: Marital status (Statut matrimonial)
- PROV: Province
- CFATINC: CF - After-tax income (Revenu familial après taxe)

Cette étape est déjà faite pour vous ici.

La base de données se nomme CIS-72M0003-E-2015\_F1.csv Le fichier CIS-72M0003-E-2015\_F1.pdf est ce qu'on appelle le dictionnaire de la base de données. Il décrit l'ensemble des variables et présente les fréquences.

## Question centrale

- Est-ce qu'il existe une inégalité de revenu entre les provinces du Canada?
- Est-ce qu'il existe une inégalité de revenu entre les hommes et les femmes au Canada?

## Comprendre notre base de données

Les différentes étapes à suivre:

1. Sélectionner les variables qui nous intéressent
2. Recoder, créer de nouvelles variables
3. Calculer les fréquences
4. Calculer les paramètres de tendance centrale
5. Calculer les paramètres de variation
6. Visualisation

Type de variables	Une seule variable
Qualitative	Diagramme de barre (diagramme en bâton)
	Diagramme circulaire
	Carte ( <b>map</b> )
Quantitative	Histogramme ( <b>geom_histogram</b> )
	Diagramme de quartile (boîte à moustaches)