



# LADY GÉRALDINE POLO ALARCON

## DATA SCIENTIST / DATA ANALYST

Stage (Fin d'études) – Disponibilité : immédiate (4 à 6 mois)

### PROFIL PERSONNEL

Étudiante en Master MAS parcours Science des données, Prédiction et Prévision économiques. Mes connaissances statistiques et ma curiosité intellectuelle sont à votre disposition. Mon objectif est d'exprimer mon plus grand potentiel et apporter une valeur ajoutée à l'équipe.

### LANGUES



Français C1



Espagnol Bilingue



Anglais C1

### PARCOURS SCOLAIRE

**Master Mathématiques Appliquées Statistiques (MAS)** 2019 – 2021

Université de Rennes 1 – Université Rennes 2  
[Plaquette MAS](#)

**Licence Mathématiques et Informatique Appliquées aux Sciences Humaines et Sociales (MIASHS)** 2016 – 2019

Université de Rennes 1  
[Plaquette MIASHS](#)

**Licence Sciences Statistiques** 2006 – 2012  
Université Centrale du Venezuela

### COORDONNEES

279A rue de Fougères

35700 Rennes

07 66 20 99 21

[lady.polo@etudiant.univ-rennes1.fr](mailto:lady.polo@etudiant.univ-rennes1.fr)

### EXPERIENCES ASSOCIATIVES



Coupe du Monde Féminine de la FIFA – France 2019



Secours Populaire 35

### COMPÉTENCES

**SGBD** : Oracle, SQL, SQLite

**DataViz** : Rmarkdown, RShiny, Bokeh

**Traitement et analyses des données** : var. quantitatives (rég. Linéaire multiple, acp), var. qualitatives (modèles binaires / multinomiaux, cluster), classification supervisée et non supervisée, analyse de séries temporelles, imputation des données

**Machine learning** : forêts aléatoires, gradient boosting, réseaux de neurones, KNN, SVM, CNN

**Techniques de sondage** : échantillonnage

### INFORMATIQUE

**SAS** (SAS SCYP Program en cours)

**Python** : Bokeh, matplotlib, numpy, pandas, scikit-learn, pytorch, keras/tensorflow, lightGBM/xgBoost, BeautifulSoup, word2vec, scrapy

**R** : Rshiny, tidyverse, ggplot2, rcharts, dataTable, caret, FactoMineR, xgboost, glm, gbm, pROC, ranger, randomForest, tensorflow

**Eviews, WinRats, SPSS IBM Office 2010, Power BI**

### RÉALISATIONS

#### Data visualisation

- Création d'une application RShiny de visualisation des logements d'AirBnb à Paris qui montre la localisation et distance par rapport aux monuments, en utilisant le langage R  
- Création d'une application Bokeh de visualisation des locaux vélos à Rennes, en utilisant Python

#### Réduction de dimensions

Analyse de Composantes Principales de délits en France 2016, en utilisant SAS

#### Classification

Détermination des niveaux d'objectivité dans les articles sportifs par l'analyse des structures grammaticales, élaboration des forêts aléatoires en utilisant le langage R

#### Séries Temporelles

Création d'un modèle de lissage Holt Winters et un modèle ARMA du prix de soja

#### Machine Learning

Application des méthodes comme KNN, Régression pénalisée, Random Forest et réseaux neurones sur le modèle AROME de Météo France (2008) pour corriger le biais de températures à 2m.

#### Text Mining

Application sur une enquête de satisfaction pour modéliser les sentiments positifs et négatifs

#### Big Data

Classement des images (pièces défectueuses ou conformes). Challenge Data de l'ENS

### PARCOURS PROFESSIONNEL

**Statisticienne, Banesco Seguros**

2013 – 2015

- Développement des indicateurs de la Direction Nationale de Souscription
- Création et présentation des dossiers des CA de la Gestion Patrimoniale
- Création et application de Sondages de Satisfaction

**Assistante de recherche, IIES Rodolfo Quintero**

Janvier – Décembre 2012

- Traiter des données et détecter les anomalies des données
- Trouver les informations de qualité pour dérouler un projet
- Rédiger les résultats et rédactions des dossiers