

# Expert ou machine ? Les deux !

scorelab

Votre expert en **Data Science** et **Machine Learning**

Finance, industrie, banque, santé, e-commerce, transport...  
Libérez la valeur de vos données grâce à nos solutions sur mesure.

# Who we are

scorelab



Guillaume Forcade

CEO & Co-founder



Jean-Baptiste  
Pautrizel

PhD Physics &  
Co-founder



Adrien Todeschini

PhD Stat ML & Chief  
Data scientist



Kévin Baudin

MSc. Stat Finance &  
Data scientist



Wassek Al Chahid

Full stack developer

# What we do

scorelab



Credit scoring

Recommender systems

Betting fraud detection

Traffic prediction

Advertising targeting

Genomics meta-analysis



Buy/Sell signals

Algo trading

Technical/fundamental analysis

Commodities



GlobalWineScore.com

Noodle.vote

NIPS / ICML

# What we use

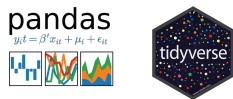
scorelab

## Data Engineering

Data collection and storage



Data transformation



Data visualization



## Machine Learning and Statistics

Supervised learning: classifiers, regressors

Unsupervised learning: dimensionality reduction, clustering

Trees and ensemble methods

Topic modeling

Latent factor models

Matrix factorization

Bayesian inference

Probabilistic programming

Non parametric statistics

Missing data imputation

Decision theory

Monte-Carlo simulation

Score aggregation



## Programming

Languages and tools



Web development



Cloud computing



Search a wine

Rankings

Critics

Primeurs

Blog

Login

English 한국어



One single score, aggregated from critics.

COLOR: Red

SCORE GWS 70 - 100

COUNTRY (OPTIONAL): Choose a country

SHOW WINE RATINGS

## LATEST GWS RANKINGS

Global Wine Score combines over 350,000 professional ratings into more than 26,000 Global Wine Scores for wines around the world. Use Global Wine Score to search for your favorite wines or discover the best wines from new regions.

### TOP AUSTRALIAN REDS 2013

99.82	1. Torbreck, Runrig Shiraz - Viognier, Barossa Valley	96.64
98.21	2. Vasse Felix, Tom Cullity, Margaret River	95.69
97.34	3. Hentley Farm, Clos Otto Shiraz, Barossa Valley	95.10
97.13	4. Clonakilla, Shiraz - Viognier, Canberra District	95.01
	5. Torbreck, The Factor Shiraz, Barossa Valley	94.58

### TOP RHONE REDS 2015

1. E. Guigal, La Turque, Cote Rotie	98.51
2. Paul Jaboulet Aine, La Chapelle, Hermitage	98.48
3. E. Guigal, La Mouline, Cote Rotie	98.40
4. Domaine Jamet, Cote Brune, Cote Rotie	98.30
5. Domaine Auguste Clape, Cornas	98.20

### TOP ARGENTINE REDS 2012

1. Bodega Aleanna, Gran Enemigo Gualtallary Cabernet Franc, Tupungato
2. Catena Zapata, Malbec Argentino, Valle De Uco
3. Bodega Aleanna, Gran Enemigo', Mendoza
4. Catena Zapata, Nicolas Catena Zapata', Mendoza
5. Terrazas De Los Andes, 'Cheval Des Andes'

NAME  
Chateau Beau-Sejour Becot, Saint Emilion Grand Cru, 2015

COUNTRY  
France

APPELLATION  
Saint Emilion Grand Cru

COLOR  
Red

GLOBAL WINE SCORE

93.68 /100 


Updated: April 1, 2018

Share 

GWS CALCULATION BASED ON

19 Wine Critics

CONFIDENCE INDEX

A+ 

scorelab

### STATISTICS

RANKING Saint Emilion Grand Cru, 2015

16. Chateau Canon La Gaffeliere, Saint Emilion Grand Cru 94.08

17. Chateau Beau-Sejour Becot, Saint Emilion Grand Cru 93.68

18. Chateau Pavie Macquin, Saint Emilion Grand Cru 93.56

19. Chateau La Gaffeliere, Saint Emilion Grand Cru 93.54

20. Chateau Trottevieille, Saint Emilion Grand Cru 93.49

21. Chateau Pavie-Decesse, Saint Emilion Grand Cru 93.48

22. Chateau Ausone, La Chapelle D'Ausone, Saint Emilion Grand Cru 93.17

23. Chateau Bellevue Mondotte, Saint Emilion Grand Cru 92.83

cellar TRACKER

CT 93.6

11 user reviews 

SCORE BY VINTAGE

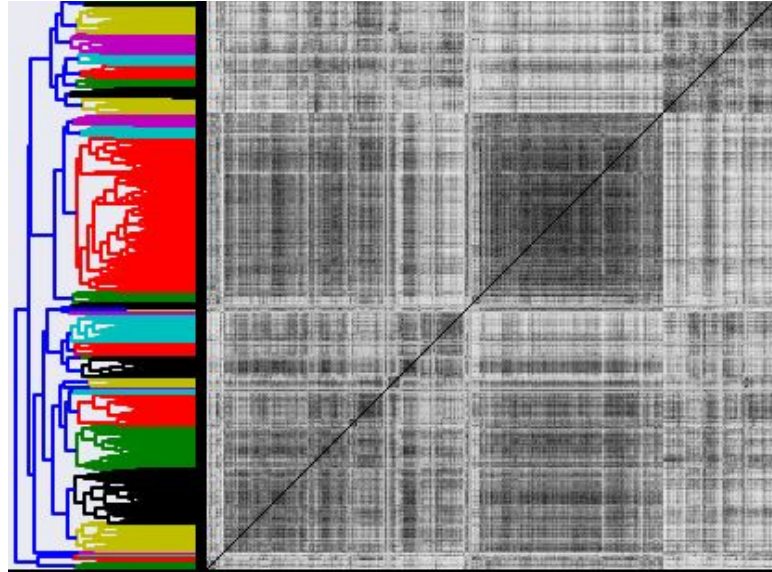


GWS EVOLUTION



# Genomics Meta-analysis

scorelab



← Back to search history

Home > Bourgogne > Morgon > Domaine Louis Tête - Morgon Les Charmeuses 2016



## DOMAINE LOUIS TÊTE - MORGON LES CHARMEUSES 2016

Domaine Louis Tête - Morgon - Red

Description: "Minéral, Boisé, Fruité"

8.2 € VAT incl.

Quantity  75 cl bottles

Add to cart



### Inspired by your browsing history

**Château Cantemerle - Les Allées de Cantemerle 2011**  
Haut-Médoc  
13.6 €

**Château de Briante 2015**  
Brouilly  
7.5 €

**Les Vignerons de Tautavel - Le Cirque - Rouge 2014**  
IGP Côtes Catalanes  
6.0 €

**M. Chapoutier - Les Meysonniers 2015**  
Crozes-Hermitage  
14.65 €

**Château de Lastours - Arnaud de Berre 2013**  
Corbières  
9.0 €

### You viewed

- **Domaine Louis Tête - Morgon Les Charmeuses 2016** Morgon 8 €
- **Charles Joguet - Cuvée de la Cure 2011** Chinon 16 €
- **Domaine Louis Tête - Côte de Brouilly 2014** Côte de Brouilly 8 €
- **La Fleur de Boüard 2011** Lalande de Pomerol 28 €<sup>8</sup> €

### Cart

- **Charles Joguet - Cuvée de la Cure 2011** Chinon 16 €
- **Domaine Louis Tête - Côte de Brouilly 2014** Côte de Brouilly 8 €
- **Domaine Louis Tête - Morgon Les Charmeuses 2016** Morgon 8 €

### Past orders



ID: 66675 35 y/o 13 orders

- **Domaine Roque Sestiere - A l'Oree des Pins 2014** Corbières 7 €<sup>7</sup> €
- **Trapiche Melodias 2015** Mendoza 10 €
- **G de Chateau Guiraud 2016 BIO** Bordeaux Sec 7 €
- **Boutinot - Domaine de Vedilhan 2015** Vin de Pays d'Oc 7 €
- **Boutinot - La Lisse Soie d'Ivoire 2015** Vin de Pays de l'Aude 8 €<sup>7</sup> €
- **Neipperg Collection Rouge 2010** Bordeaux 5 €
- **Les Jamelles - Merlot 2015** Vin de Pays d'Oc 6 €
- **Chateau Coudray Montpensier - Tradition 2015** Chinon 7 €
- **Maison Tardieu Laurent - Les Grandes Bastides Cotes du Rhone rouge 2014** Cotes-du-Rhone 8 €
- **Decignano Dei Barbi - Villa Barbi Rosso 2014** Ombrie 8 €
- **Domaine Coudoulis - Evidence 2016** Lirac 8 €
- **Domaine Louis Tete - Cote de Brouilly 2014** Cote-de-Brouilly



# Nips17.ml

scorelab

### Register

Which topics are you interested in?

- ✓ Optimization
- ✓ Theory
- ✓ Deep Learning
- ✓ RNNs
- ✓ Reinforcement Learning
- ✓ Supervised Learning
- ✓ Graphs
- ✓ Matrices

### Schedule

Mon	Tue	Wed	Thu	Fri	Sat
09:00 AM	10:20 AM	01:50 PM	02:50 PM		06:30 PM
		02:50 PM - 03:05 PM • Hall A	02:50 PM - 03:05 PM • Hall C		
		<b>ELF: An Extensive, Lightweight a...</b> Yuandong Tian, Qucheng Gong, W...	<b>The Marginal Value of Adaptive G...</b> Ashia C Wilson, Becca Roelofs, Mit...		
		03:05 PM - 03:20 PM • Hall A	03:05 PM - 03:20 PM • Hall C		
		<b>Imagination-Augmented Agents f...</b> Seb Racanière, Theophane Weber,...	<b>Can Decentralized Algorithms Ou...</b> Xiangru Lian, Ce Zhang, Huan Zha...		
		03:20 PM - 03:50 PM • Hall A	03:20 PM - 03:50 PM • Hall C		
		<b>Reinforcement Learning, Deep Le...</b>	<b>Optimization</b>		

### Library

Algo Representation

**Hunt For The Unique, Stable, Sparse And Fast Feature Learning On Graphs**  
Saurabh Verma, Zhi-Li Zhang

BOOKMARK LIKE

Algo CollabFiltering

**Mixture-Rank Matrix Approximation for Collaborative Filtering**  
Dongsheng Li, Kehan Chen, Wei Liu, Tun Lu, Ning Gu, Stephen Chu

BOOKMARK LIKE

Applied RecSys

**Scalable Demand-Aware Recommendation**

### Explore

Theory Complexity RECOMMENDED

**Near Optimal Sketching of Low-Rank Tensor Regression**  
Xingguo Li, Jarvis Haupt, David Woodruff

DISMISS BOOKMARK LIKE

Algo Regression RECOMMENDED

**A Unified Approach to Interpreting Model Predictions**  
Scott M Lundberg, Su-In Lee

DISMISS BOOKMARK LIKE

Algo Sparsity CompressSensing RECOMMENDED

**Parametric Simplex Method for Sparse Learning**  
Haotian Pang, Han Liu, Robert J Vanderbei, Tuo Zhao

DISMISS BOOKMARK LIKE

# Doodle for Notation

Evaluate and rate any contest in an easy manner.  
In a few simple steps create your **Pitch Contest.**

Create a contest

## How it works



### 1. Create a contest

In 2 clicks, create a secured contest and set your preferences.



### 2. Define items and criteria

Declare unlimited candidates or items and their evaluation criteria.



### 3. Invite participants to vote

Share your Noodle with your friends, colleagues or strangers!



### 4. And the winner is...

Once the votes are in, see the result! Share the outcome on social networks.



## Overview

Welcome, Adrien

Evaluate all the items of *Soutenances de stage 2018*. If you prefer you can instead evaluate by criteria.

ITEMS CRITERIA

- Yang >
- Kevin >
- Lauren >
- Wassek >
- Mirko >
- Phi >

Edit the contest



Lauren

Oral

4

Diapos

4

Respect du temps

5

Questions

3

Rapport

5

Good job!

Rate next item

Back to overview



## Soutenances de stage 2018 results

1 voter | created 2018-12-04

Share

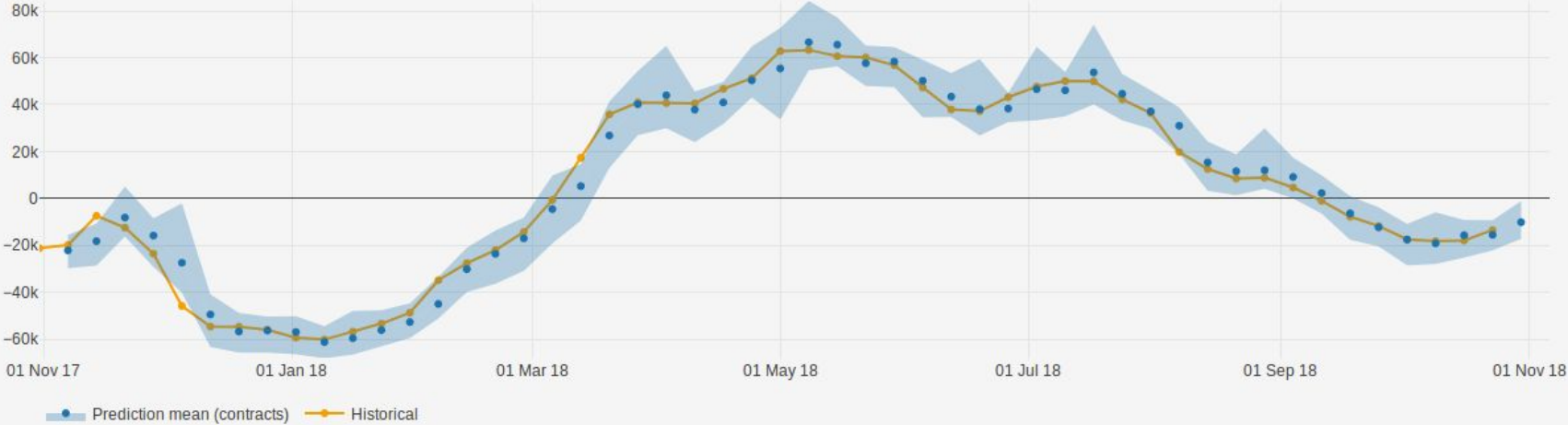
Name	Rank	Score
Lauren	1	4.33
Kevin	2	4.0
Wassek	3	4.0
Phi	4	3.83
Mirko	5	3.5
Yang	6	3.0

Back to overview

# Market predictions

## Managed Money Positions Net (Europe)

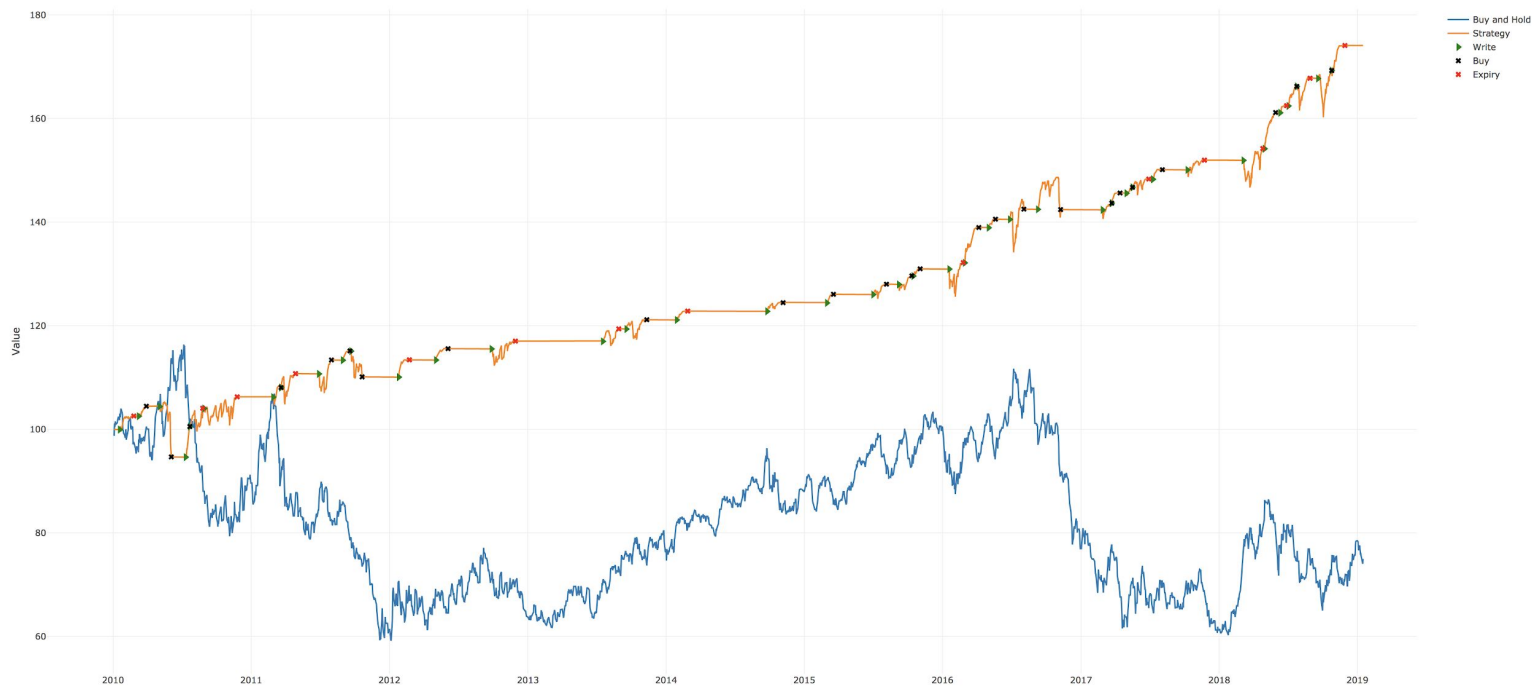
Managed Money Positions Net (EU)



# Algo Trading

scorelab

Option Strategy - Based on Reversion signal



Preacor.fr

# Votre accord en un temps record

Renseignez votre situation en seulement trois minutes afin de calculer votre pourcentage de chances d'obtenir un prêt immobilier.

[Commencer la simulation >](#)

Vous avez déjà un compte ? [Connectez-vous](#)



- ✓ État civil
- ✓ Revenus
- ✓ Comportement bancaire
- ✓ Charges & Patrimoine
- ✓ Plan de financement
- **Résultat**

## Résultats



**Vous avez obtenu un score de 83%**

Votre projet ne semble présenter aucune difficulté particulière.

## Mon projet immobilier

Prix du bien	<input type="text" value="30000"/>	€	Frais de notaire	<input type="text" value="22500"/>	€
Prix des travaux	<input type="text" value="60000"/>	€	Frais d'agence	<input type="text" value="0"/>	€
Frais divers	<input type="text" value="0"/>	€			
<b>Prix du projet immobilier</b>	<b><input type="text" value="382500"/></b>	<b>€</b>			



## Etat Civil

### Projet

Résidence principale :

Résidence secondaire :

Achat locatif :

Mixte: locatif et principal :

Autre :

### Actuellement vous êtes

Propriétaire de votre résidence principale :

Locataire :

Hébergé à titre gratuit :

Logement de fonction :

Autre :

## scorelab

### Pourcentage d'endettement

0 - 20 % :

20 - 25 % :

25 - 30 % :

30 - 33 % :

33 - 38 % :

38-50 :

50-100 :

>100 :

# Data

scorelab

~8 ans d'historique, ~35000 dossiers, ~70 colonnes

## Dossier

numero\_dossier  
date\_creation  
etape\_dossier  
nb\_emp

## Revenus

revenu\_alloc  
revenu\_mensuel  
catpro  
type\_contrat  
employeur  
anciennete\_job

## Crédits

nb\_prets\_immo  
crd\_prets\_immo  
mensu\_prets\_immo  
nb\_prets\_conso  
crd\_prets\_conso  
mensu\_prets\_conso

## Charges

charge\_pension  
charge\_impots  
charge\_loyer  
charge\_garde\_enfants  
charge\_autre  
charge\_autre\_nature

## (Co)Emprunteur

cp, ville  
naissance  
isproprio, isloc, isheberge  
situ  
nbenfants  
banque  
anciennete\_banque

## Patrimoine

mnt\_epargne  
nb\_biens  
valeur\_biens  
revenu\_biens

## Projet

cp\_bien\_projet  
ville\_bien\_projet  
surface\_bien\_projet  
terrain\_bien\_projet  
usage\_projet  
isprimoaccedant

## Financement

mnt\_terrain  
mnt\_logement  
mnt\_travaux  
mnt\_frais\_agence  
mnt\_frais\_notaire  
honos\_courtier  
mnt\_apport  
mnt\_finance  
banque\_pret  
nb\_prets\_differe\_partiel  
duree\_mois\_financement  
refus\_motif

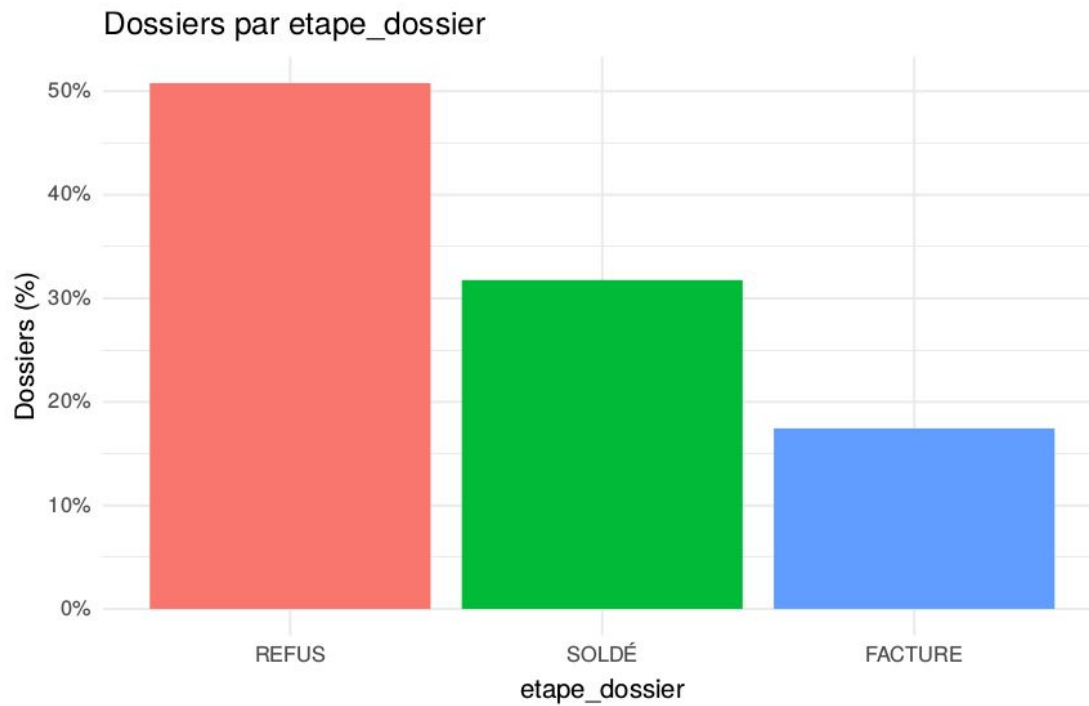
# Variables absentes

scorelab

- Type profession libérale
- Comportement bancaire : solde moyen, découverts
- Taille de l'entreprise
- Liasses fiscales entrepreneurs
- Faire appel à un constructeur
- Assurance dommage ouvrage

# Variable à prédire

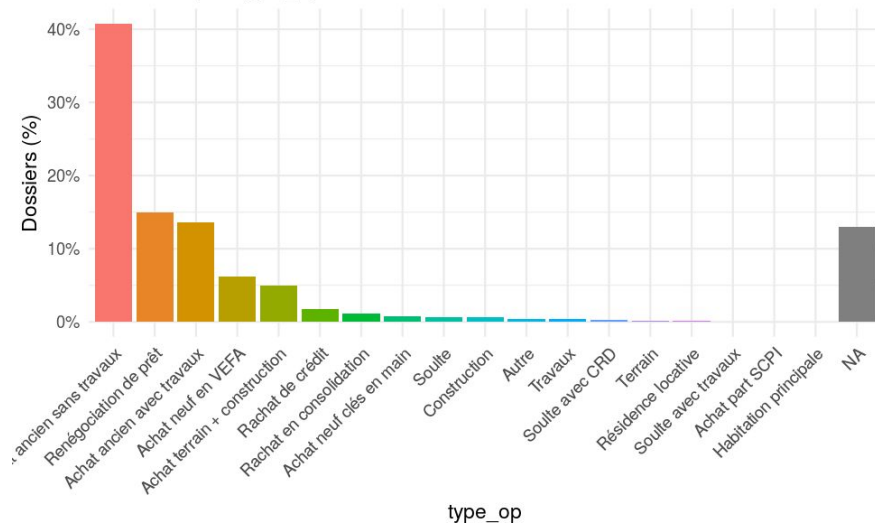
scorelab



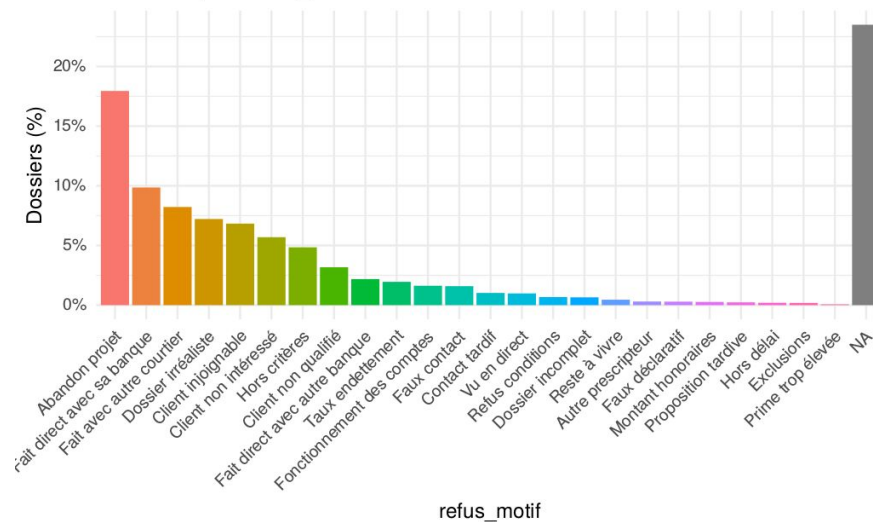
# Filtrage des cas pertinents

- type\_op = Achat ancien avec travaux, Achat ancien sans travaux, Achat neuf clés en main, Achat neuf en VEFA, Achat terrain + construction
- refus\_motif = Fonctionnement des comptes, Taux endettement, Reste à vivre, Hors critères, Prime trop élevée
- On obtient **~13000 dossiers**, **~8% de refus** (Déséquilibre)

Dossiers par type\_op

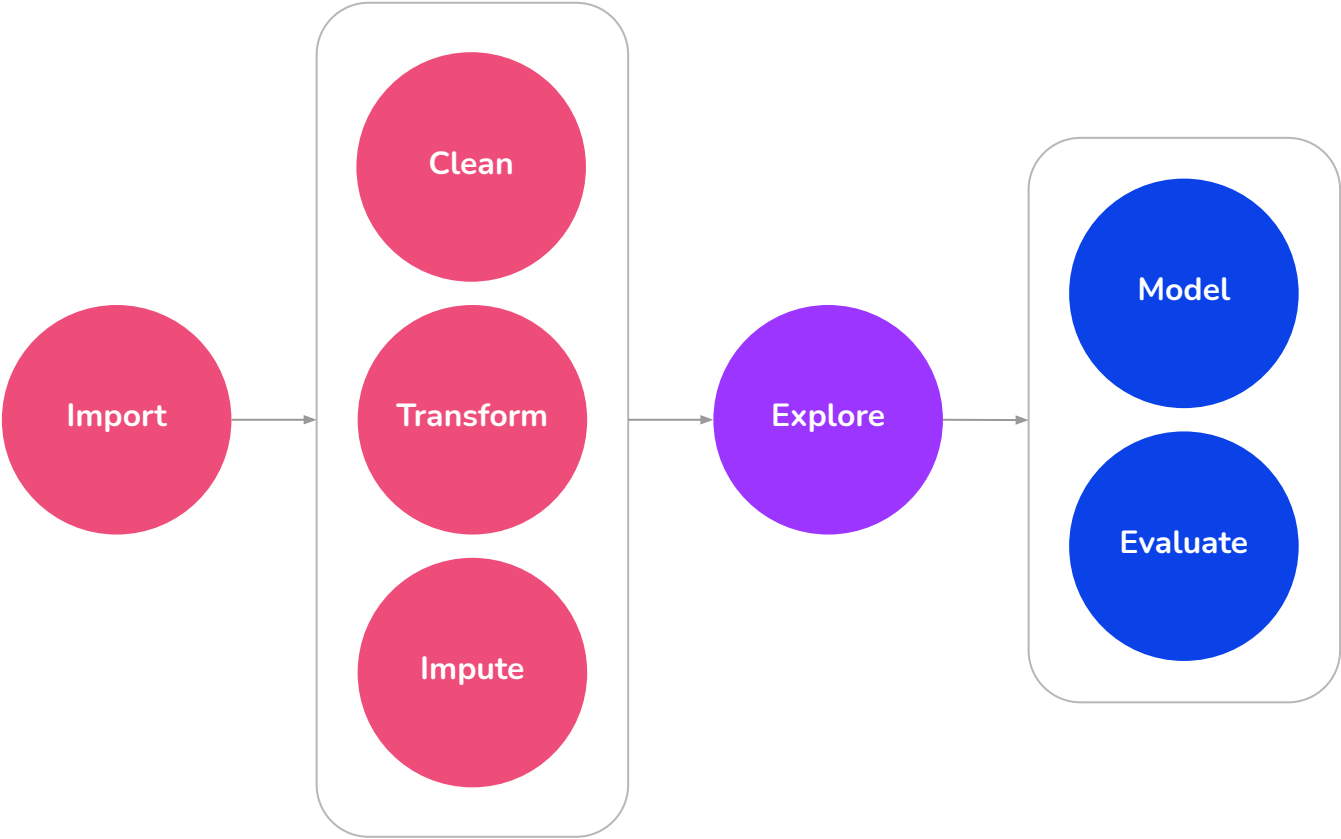


Dossiers par refus\_motif



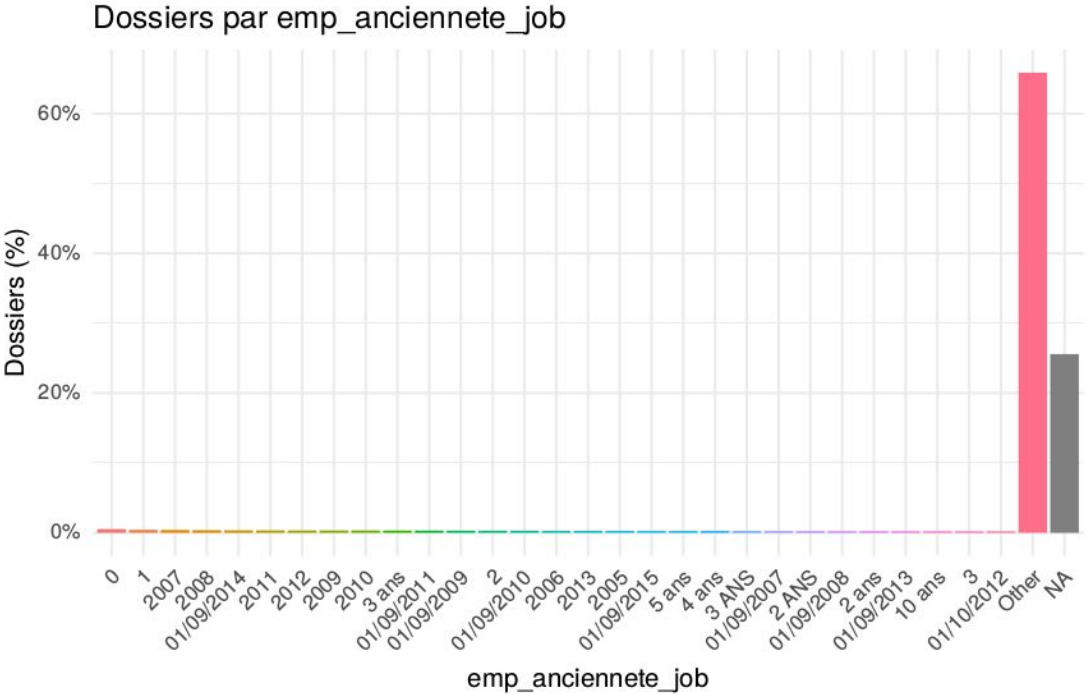
# Pipeline

scorelab



# Nettoyage

scorelab

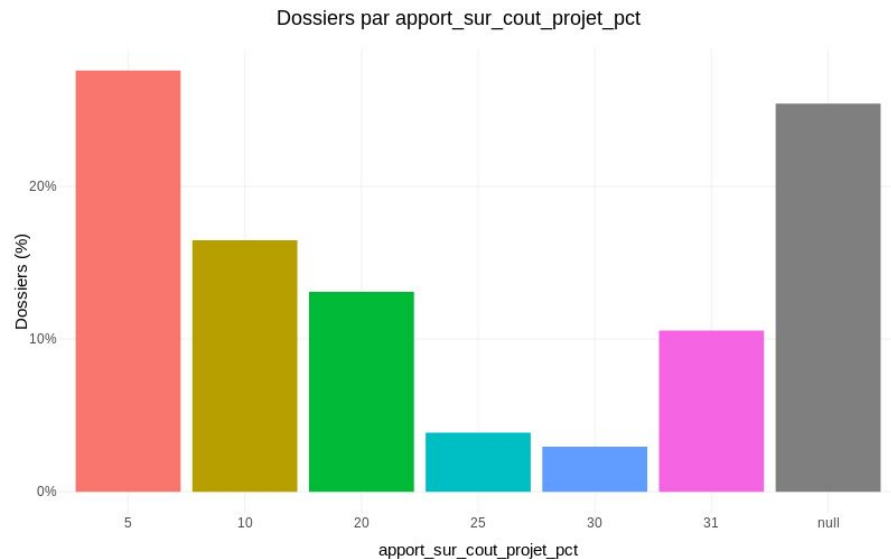


# Transformation

scorelab

## Feature engineering

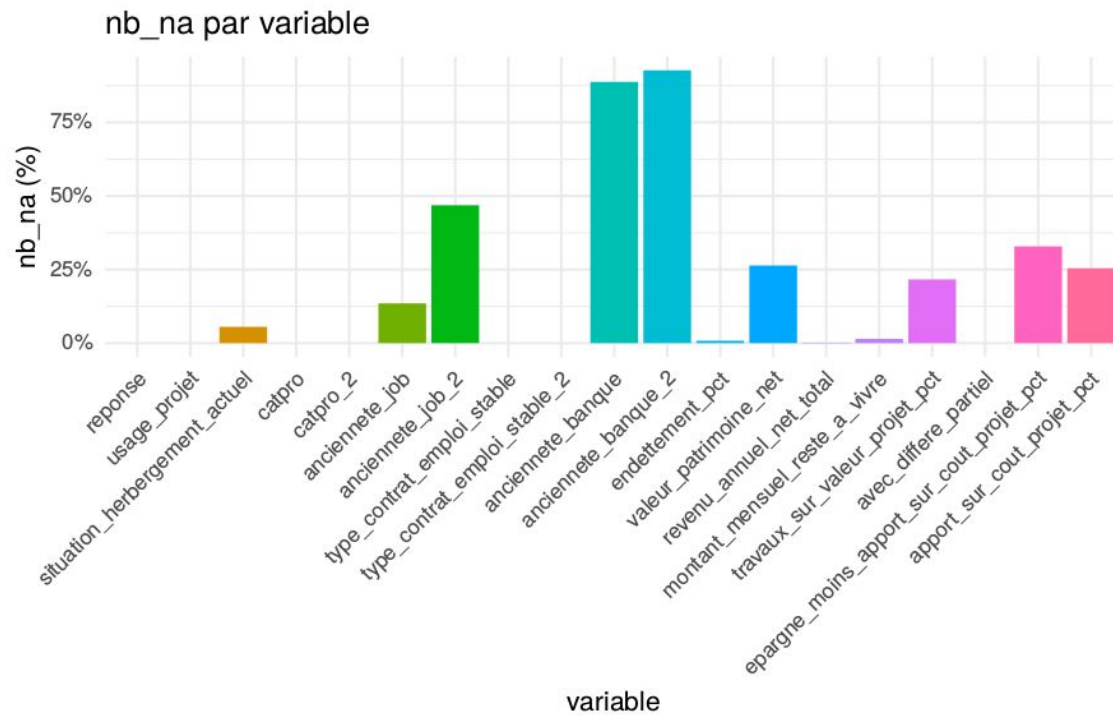
- Pourcentage d'endettement
- Reste à vivre par personne
- Valeur patrimoine net
- Revenu annuel net total
- Pourcentage des travaux sur prix du bien
- Pourcentage d'apport sur coût du projet
- Epargne après opération sur coût du projet





# Imputation

scorelab



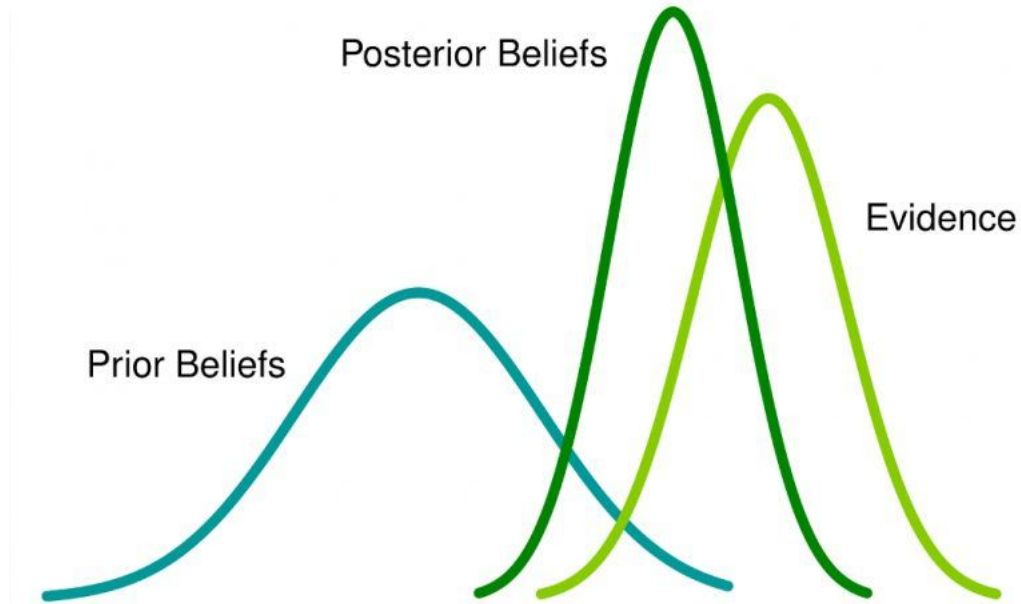
# Problèmes rencontrés

scorelab

- Dossiers non pertinents → écartés
- Données aberrantes → seuillées
- Données non typées, mal encodées → parsing
- Données manquantes → imputation
- Classes déséquilibrées → oversampling
- Variables absentes ou sous-représentées → modèle bayésien

# Inférence bayésienne

scorelab



# Régression logistique bayésienne

scorelab

Vraisemblance

$$p(Y|\beta, X)$$

$$y_i|p_i \sim \text{Ber}(p_i)$$

$$\log\left(\frac{p_i}{1-p_i}\right) = \frac{1}{\sigma_i} \left( \beta_0 + \sum_{j=1}^p \beta_j x_{ij} \right)$$

A priori

$$p(\beta)$$

$$\beta_j \sim \mathcal{N}(\mu_j, \sigma_{c_j}^2)$$

A posteriori

$$p(\beta|Y, X) \propto p(Y|\beta, X)p(\beta)$$

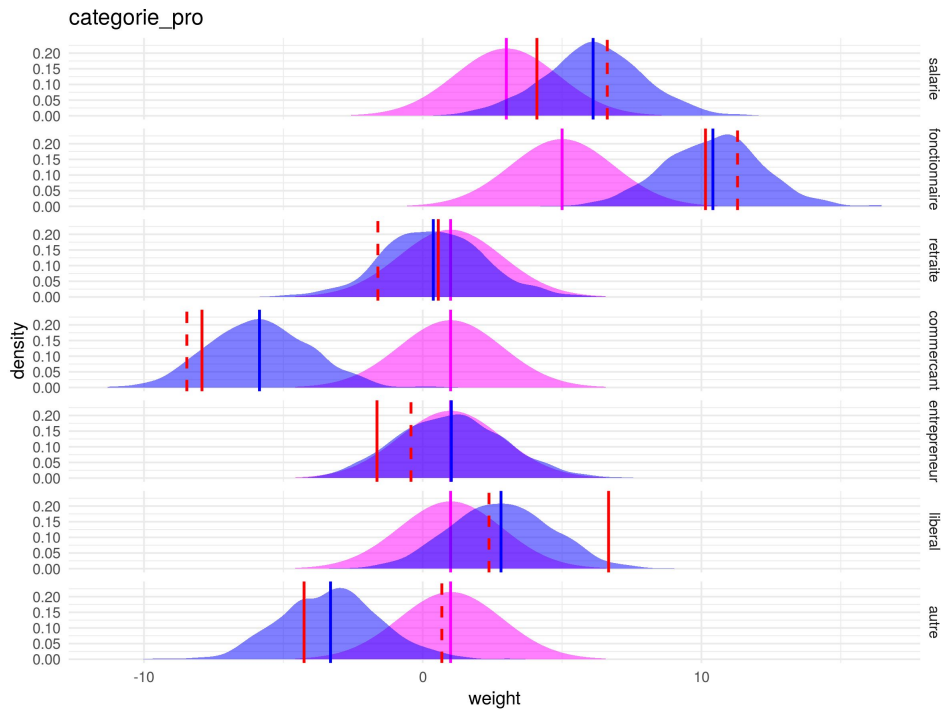
# JAGS code

scorelab

```
model {  
  
  # likelihood  
  for (i in 1:n) {  
    y[i] ~ dbern(prob[i])  
  
    logit(prob[i]) <- ( beta_0 + inprod(X[i,], beta) + inprod(X_coemp[i,], beta[1:p_coemp] ) / scale[i]  
  }  
  
  # prior  
  beta_0 ~ dnorm(0, 1/sigma^2)  
  
  for (k in 1:n_var) {  
    for (j in from[v]:to[v]) {  
      beta[j] ~ dnorm(mu[j], 1/(sigma_var[k])^2)  
    }  
  }  
}
```

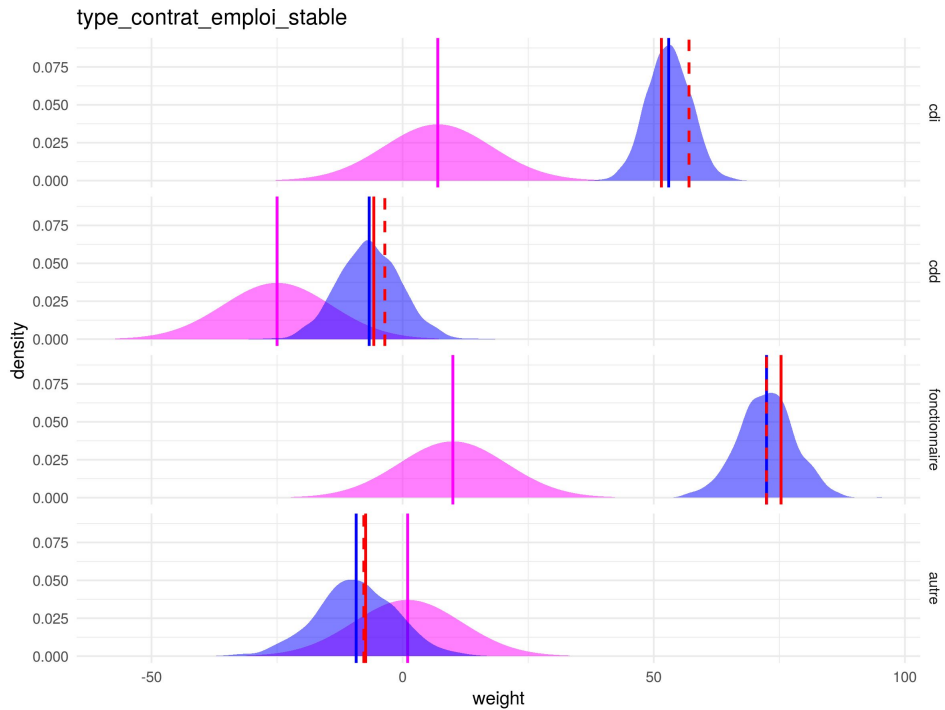
# Coefficients

scorelab



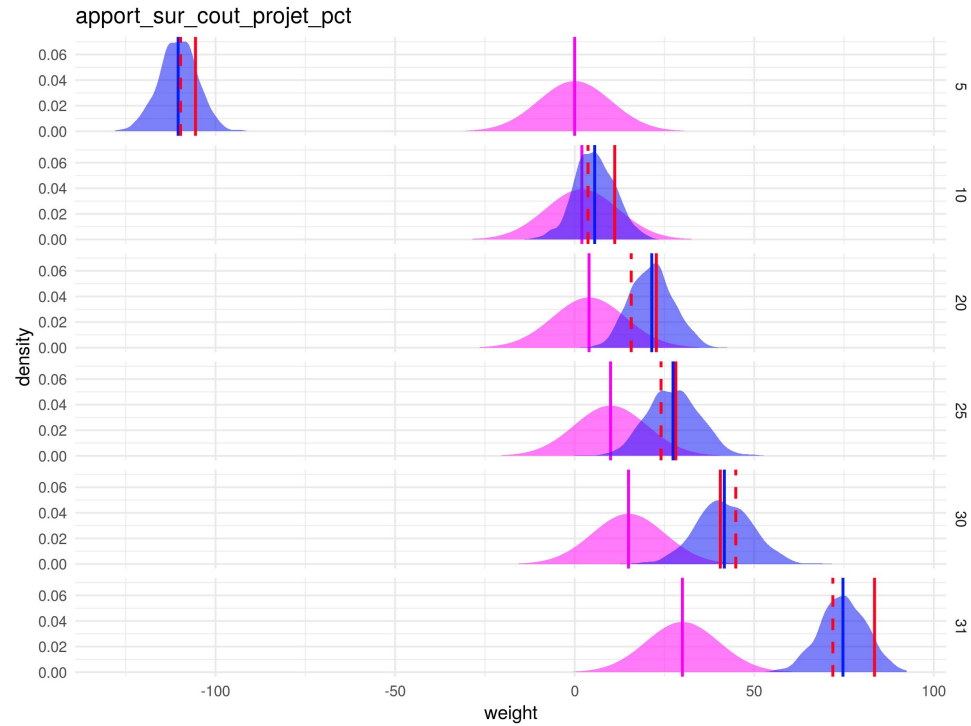
# Coefficients

scorelab



# Coefficients

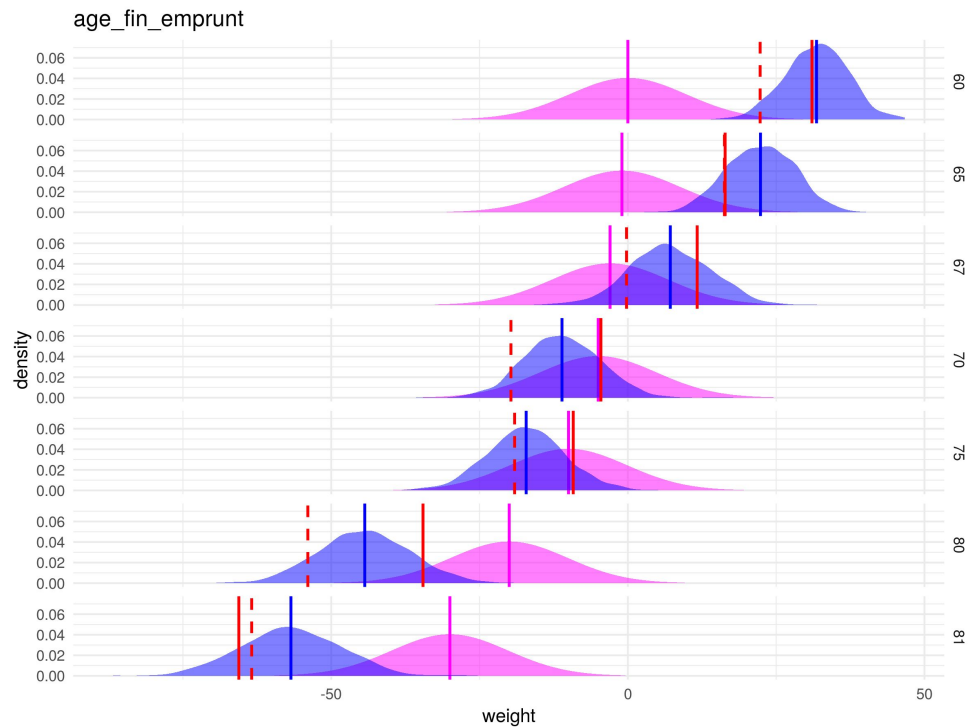
scorelab





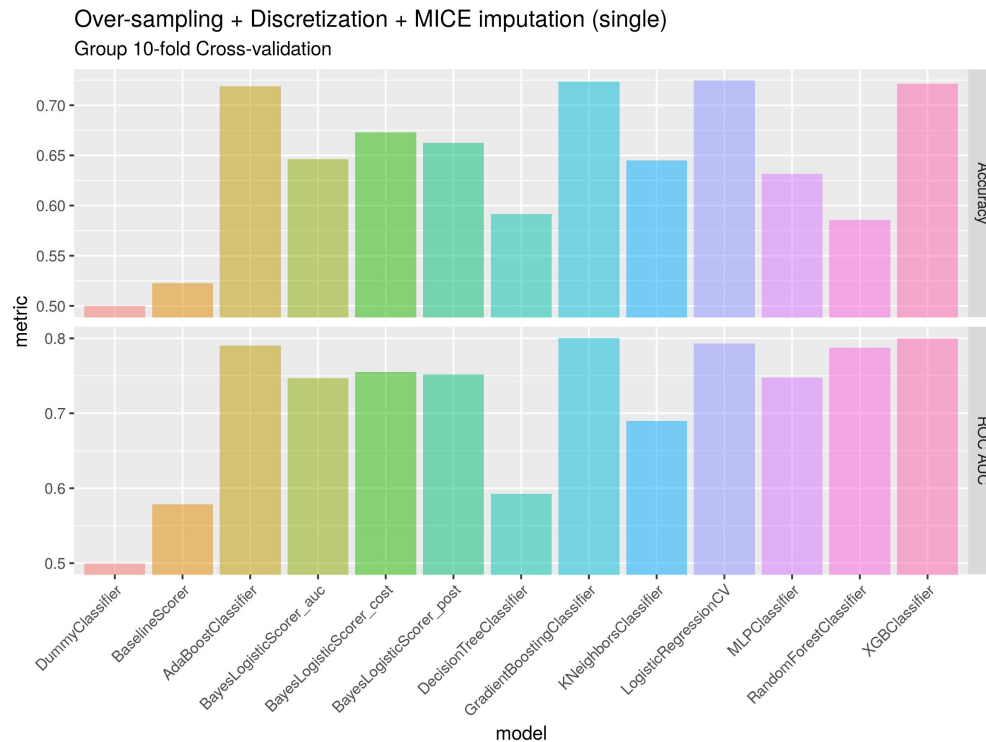
# Coefficients

scorelab



# Comparison

scorelab





contact@scorelab.io

 @scorelab\_io

