**Cumulative Accuracy Profile (CAP)**

Penons le modèle final de la partie `Robust Geodemographic Segmentation Model`

Une image contenant texte, capture d’écran, Police, document

Description générée automatiquement

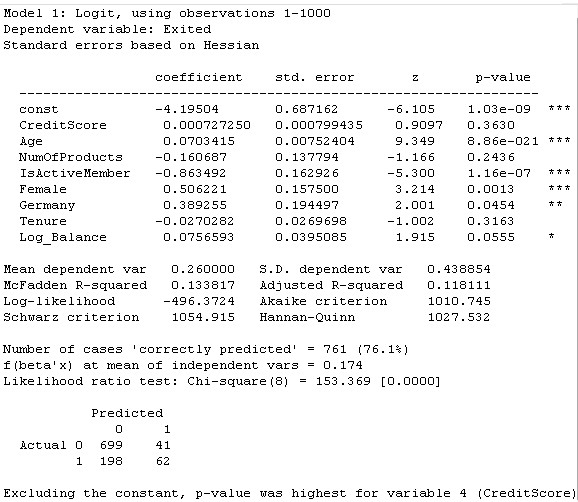
Nous avons créé les fichiers des prédictions `Churn-Modelling\_Forecast.csv`, puis nous avons travaillé avec le fichier Excel(Churn-Modelling\_Forecast.xlsx) pour obtenir le CAP.

Ensuite, Nous avons travaillé avec le template donné qu’avec les 1000 premières valeurs ordonnées (voir CAP-Curve-Template\_1.xlsx)

**Training vs Testing set**

Le training set représente 80% de nos dataset, il permet de construire le modèle prédictif et 20% restant est le testing set qui permet donc de tester ce modèle.

Application sur les données de `Churn-Modelling`, nous avons le training set (P12-Churn-Modelling.csv de 1 a 10000 valeurs) et le testing set (P12-Churn-Modelling-Test-Data de 10001 a 11000 valeurs). On considère toujours le modèle final de la partie `Robust Geodemographic Segmentation Model` que nous appliquons a P12-Churn-Modelling-Test-Data. Nous trouvons :



Les valeurs prédites sont enregistrées dans tab\_predict\_TestData.pdf.

Nous allons comparer les valeurs prédites du testing set avec les valeurs réelles de Exited qui sont dans le fichier P12-Churn-Modelling-Test-Data.csv. Nous allons enregistre cette comparaison dans P12-Churn-Modelling-Test-Data\_forecast.csv.

Afin d’évaluer la performance des prédictions sur TestData, nous allons range par ordre décroissant la colonne Proba\_hat\_TestData et réutilise le template (voir CAP-Curve-Template-Test-Data). Nous trouvons ce CAP suivant :

