

Plan de Préparation Data Analyst - Examen UniBank Haiti

Vue d'ensemble

Contexte: Préparation pour un poste de Data Analyst dans une banque commerciale

Institution: UniBank Haiti

Format: Questions d'entretien + Études de cas pratiques

Niveau minimum: Intermédiaire

Structure de la Préparation

#	Document	Sujet	Priorité
1	01_Data_Visualization.md	Visualisation de données - Statistiques par graphe	Haute
2	02_EDA_Data_Wrangling.md	Exploration et nettoyage des données	Critique
3	03_Indicateurs_Indices.md	Indicateurs et indices (théorie + pratique)	Haute
4	04_Statistiques_Descriptives.md	Statistiques descriptives	Critique
5	05_Statistiques_Inferentielles.md	Échantillonnage et tests d'hypothèses	Haute
6	06_Analyse_Univariee_Multivariée.md	Analyse uni/multivariée	Haute
7	07_Probabilites.md	Concepts clés de probabilité	Moyenne
8	08_Business_Intelligence_Bancaires.md	Banques et modèles bancaires	Critique
10	10_SQL_Data_Analyst.md	SQL avancé pour analysts	Haute
11	11_Data_Engineering_Fundamentaux.md	Éléments fondamentaux Data Engineering	Moyenne
12	12_Types_Variables.md	Tous les types de variables	Critique
13	13_Types_Modeles.md	Types de modèles (général + bancaire)	Haute
14	14_Machine_Learning.md	ML fondamental + cas bancaires	Haute
15	15_Regression_Lineaire.md	Régression simple et multiple, diagnostics	Critique
16	16_Series_Temporelles.md	Prévisions et modèles temporels	Haute
17	17_Tests_Non_Parametriques.md	Tests alternatifs (Mann-Whitney, etc.)	Haute

#	Document	Sujet	Priorité
18	18_AB_Testing_Experimentation.md	A/B Testing et design expérimental	Critique
19	19_Ethique_Gouvernance_Donnees.md	Éthique, biais, gouvernance des données	Critique

Dossiers Complémentaires

/revisions/

- Manuel_Revision_Complet.md - Révision finale
- Fiches_Synthese.md - Mémos ultra-rapides avant examen
- Manuel_Revision_Cas_Speciaux.md - Valeurs manquantes, outliers, ACP, ANOVA
- Mnemoniques_Examens.md - Mnémotechniques pour mémorisation

/tests/

- ctx_global/ - 4 tests globaux (30-50 questions)
- ctx_*/ - Tests individuels par sujet (2 par sujet)
- ctx_regression/ - Tests régression linéaire (2)
- ctx_series_temp/ - Tests séries temporelles (2)
- ctx_non_param/ - Tests non-paramétriques (2)
- ctx_ab_testing/ - Tests A/B testing (2)
- ctx_cas_speciaux/ - Tests cas spéciaux (2)

/FULL/

- Études de cas complètes par domaine

/out_of_scope/

- Documents hors périmètre de l'examen (PowerBI, DAX)
-

Principe Pareto Appliqué (20/80)

Concepts à priorité absolue pour Data Analyst Bancaire

Types de Variables: - Variables qualitatives (nominales, ordinaires) - Variables quantitatives (discrètes, continues) - Niveaux de mesure et implications pour l'analyse

Visualisation de données: - Histogrammes, Box plots, Scatter plots - Choix du graphe selon le type de données - Interprétation statistique

EDA & Data Wrangling: - Détection et traitement des outliers - Gestion des valeurs manquantes - Feature engineering basique

Statistiques Descriptives: - Moyenne, médiane, mode - Variance, écart-type - Quartiles, percentiles

Statistiques Inférentielles: - Test t, Chi-carré, ANOVA - p-value et intervalles de confiance - Corrélation vs causalité

Business Intelligence Bancaire: - KPIs bancaires (NPL, LTV, CAR, ROA, ROE) - Segmentation clients - Analyse de risque crédit

Types de Modèles: - Modèles descriptifs vs prédictifs vs prescriptifs - Classification vs régression vs clustering - Modèles de risque bancaire (PD, LGD, EAD)

Machine Learning: - Algorithmes supervisés (logistique, arbres, forêts) - Algorithmes non supervisés (K-Means, détection anomalies) - Applications: scoring crédit, détection fraude, churn

SQL pour Analysts: - Window functions (ROW_NUMBER, RANK, LAG, LEAD) - CTEs et requêtes complexes - Optimisation des requêtes - N+1 query problem

Régression Linéaire: - Modèle simple et multiple ($Y = \beta_0 + \beta_1 X + \epsilon$) - Hypothèses LINE (Linéarité, Indépendance, Normalité, Égalité variances) - Diagnostics (VIF, Durbin-Watson, R^2 , résidus) - Interprétation des coefficients

Séries Temporelles: - Composantes (Tendance, Saisonnalité, Cycle, Irrégulier) - Stationnarité et tests (ADF, KPSS) - Modèles ARIMA/SARIMA, Holt-Winters - Métriques (MAPE, AIC/BIC)

Tests Non-Paramétriques: - Mann-Whitney U, Wilcoxon, Kruskal-Wallis - Corrélation de Spearman vs Pearson - Quand utiliser (non normalité, petits échantillons)

A/B Testing et Expérimentation: - Design expérimental (randomisation, stratification) - Calcul taille d'échantillon et puissance statistique - Analyse des résultats (z-test proportions, intervalles de confiance) - Pièges: peeking, multiple testing, effet nouveauté

Éthique et Gouvernance des Données: - Biais algorithmiques (Disparate Impact, proxy discrimination) - Explicabilité des modèles (SHAP, feature importance) - Protection des données (anonymisation, consentement) - Conformité réglementaire (BRH, principes FAIR)

Contexte Bancaire Haïtien - Points d'Attention

Métriques Clés UniBank

- **Portefeuille de prêts:** NPL ratio, provision coverage
- **Dépôts:** Mix des dépôts, cost of funds
- **Rentabilité:** NIM, cost-to-income ratio
- **Liquidité:** LCR, NSFR
- **Capital:** CAR, Tier 1 ratio

Réglementation

- BRH (Banque de la République d'Haïti) standards
- Reporting réglementaire
- Anti-money laundering (AML)

Défis Data Spécifiques

- Qualité des données historiques
- Intégration de sources multiples
- Reporting en temps réel vs batch

Matériel Nécessaire

Outils Techniques

- Python (Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn)
- SQL (PostgreSQL ou SQL Server)
- Excel avancé
- Outils de visualisation (Matplotlib, Seaborn)

Documentation

- Ce guide de préparation complet
 - Fiches de synthèse pour révision rapide
 - Tests d'entraînement avec corrections
-

Stratégie de Préparation Recommandée

Phase 1: Fondamentaux (Jours 1-3)

1. Types de variables et niveaux de mesure
2. Statistiques descriptives
3. Probabilités de base

Phase 2: Analyse (Jours 4-6)

1. EDA et Data Wrangling
2. Statistiques inférentielles
3. Analyse uni/multivariée

Phase 3: Modélisation (Jours 7-9)

1. Types de modèles (descriptifs, prédictifs)
2. Machine Learning fondamental
3. SQL avancé
4. Régression linéaire et diagnostics

Phase 4: Contexte Bancaire (Jours 10-11)

1. BI bancaire et KPIs
2. Modèles de risque (scoring, fraude)
3. Séries temporelles (prévision dépôts, liquidité)
4. Études de cas bancaires

Phase 5: Expérimentation et Éthique (Jours 12-13)

1. A/B Testing et design expérimental
2. Éthique et gouvernance des données
3. Explicabilité des modèles (SHAP)
4. Biais algorithmiques et équité

Phase 6: Révision (Jours 14-15)

1. Tests pratiques
 2. Fiches de synthèse
 3. Simulation d'entretien
-

Conseils pour l'Entretien

Questions Techniques

- Toujours contextualiser avec des exemples bancaires
- Montrer la compréhension business, pas juste technique
- Expliquer le “pourquoi” derrière chaque choix

Études de Cas

- Structurer la réponse: Problème → Données → Méthode → Résultat
- Mentionner les limites et hypothèses
- Proposer des next steps

Communication

- Vulgariser pour un public non-technique
- Utiliser des analogies bancaires
- Être prêt à défendre ses choix méthodologiques

Objectif: Démontrer une maîtrise technique solide avec une compréhension approfondie du contexte bancaire haïtien.