Nwajiunor Edgar-Odurombi

Hack The Feed: Insights From Social Media Data

Outline

- Introduction
- Data Exploration and Feature Engineering
- Insights

Start Page

Introduction

Context

This project will extract, clean and visualize different datasets about social media using pandas. Social media refers to the means of interactions among people in which they create, share, and/or exchange information and ideas in virtual communities and networks. Each dataset represents features one of the following social media platforms; LinkedIn, Twitter, Instagram and Twitter for Stanbic IBTC.

Goals for the EDA

To investigate:
Which content type has the highest count?

Which content type has the highest average impressions across platforms?

Which network has the highest engagement?

Which content type has the highest reactions across all platforms?

Summary of steps

The Steps taken are:

- Data Extraction
- Data Cleaning
- Data Analysis
- Data Interpretation

Data Exploration

Dataset Overview

The merged dataset has 39062 rows (for observations) and 147 columns (for variables).

The variables are of different data types, including strings, floats, and integers.

The dataset contained null values and empty columns

Data Exploration

Data Exploration

Cleaning processes such as change of datatype was carried out in Excel. The Python programming language was used to handle the EDA cleaning processes, and the steps were documented in a Python notebook. Because not all columns are required to answer this task's questions, some columns were removed. Removing these unnecessary variables will help focus on the variables that will focus on the business objective in this task.

Feature Engineering

This involved the transforming of features (variables) to carry out more analysis. The new variables include the day of the week and engagement rate.

Insights

Summary of Key Findings:

Polls as a Content Type: Polls, as a content type, have the highest average impressions, indicating that they generate a significant amount of visibility.

Facebook Engagement: Facebook leads in terms of average engagements, suggesting that the audience on this platform is more active and interactive with the content.

Photo Content and Reactions: Photo content types receive the highest number of reactions, indicating that visual content tends to resonate well with the audience.

Photos as Engaging Content: Across all social media platforms, photos emerge as the most engaging type of content. This highlights the importance of visual elements in capturing audience attention.

Insights Cont’d)

Facebook's Engagement Dominance: Facebook stands out as the most engaging platform for the client, likely due to its high user base and active community.

Engagement by Day of the Week: Wednesdays witness the highest number of engagements and shares, suggesting that mid-week posts are particularly effective in driving audience interaction.

LinkedIn's Performance: LinkedIn demonstrates the highest engagement and click-through rate among the platforms analyzed, followed closely by Twitter. This indicates that these platforms are valuable for achieving specific marketing objectives.