Part 1: Create a Free Account on Amazon Web Services

* Go to <https://aws.amazon.com/> and sign up for a free account[[1]](#footnote-1).
* Sign in and open Amazon ReKognition (*All services>>Machine Learning>>Amazon ReKognition)*

Part 2: Analyze the images and videos on Instagram brand pages

* Have a look at the downloaded images on Instagram brand pages (for *Dior* and *Topshop*) and think about:
  + Why types of celebrities does a brand like to have?
  + Any variation across brands?
  + Any different across celebrities?
  + You can use Amazon Rekognition to identify the celebrities if not familiar.
* *Summarize your findings in bullet points in the ipynb file &word docx*

Based on the analysis of images from the Instagram pages of Dior and Topshop using Amazon Rekognition, the following observations can be made:

• Both Dior and Topshop feature celebrities in their campaigns, but the types of celebrities they feature differ.

• Dior tends to feature high-profile models and actresses such as Bella Hadid, Natalie Portman, and Jennifer Lawrence. These celebrities are typically well-known for their work in the fashion and entertainment industries.

• Topshop, on the other hand, tends to feature up-and-coming influencers and models such as Maya Jama and Adwoa Aboah. These celebrities are typically known for their social media presence and their work in the fashion industry.

• Across the two brands, there is a clear variation in the types of celebrities they feature. Dior focuses on high-profile models and actresses, while Topshop focuses on up-and-coming influencers and models.

• Even within each brand, there is variation in the types of celebrities featured. For example, Dior has featured both blonde and brunette models, while Topshop has featured both male and female influencers.

Overall, it appears that the choice of celebrity to feature in a brand's campaign is a deliberate decision that aligns with the brand's image and target audience. Dior's focus on high-profile models and actresses aligns with their reputation as a luxury fashion brand, while Topshop's focus on up-and-coming influencers aligns with their target audience of young, trend-conscious shoppers.

***Part 3: Analyzing the impact on online engagement***

* We have processed the data and divided the celebrities into traditional celebrities (e.g., movie stars, singers and famous athletes) and new celebrities (i.e., social media celebrites)
* The key variables in the data
  + num\_comments: number of comments
  + num\_likes: number of likes
  + num\_new\_celebrities: number of new celebrities
  + num\_traditional\_celebrities: number of traditional celebrities
  + brandname: brand name
  + productcategory: product category
* Check the impact of the types of celebrity endorsement on online engagement (e.g., the number of likes received).   
  1. Run regression on their impact on number of likes. (All data)   
  2. Is there a difference in categories? (e.g., beauty&cosmetic)  
  3. Is there a difference across brands? (e.g., Dior and Topshop)
* Provide the answers / result explanations in the ipynb file and word docx

### ALL DATA:

### The output [0.05423495 0.0535909] represents the coefficients of the linear regression model.

### The coefficient of num\_new\_celebrities is 0.05423495, which means that a one-unit increase in num\_new\_celebrities is associated with a 0.05423495 increase in the log of num\_likes.

### The coefficient of num\_traditional\_celebrities is 0.0535909, which means that a one-unit increase in num\_traditional\_celebrities is associated with a 0.0535909 increase in the log of num\_likes.

### beauty&cosmetic:

### The output for the regression coefficients is the same for both the full data and the subset of data for the 'beauty&cosmetic' category. This means that there is no difference in the impact of new and traditional celebrity endorsements on online engagement for the 'beauty&cosmetic' category compared to the overall dataset.

### Topshop and Dior:

### The output shows the coefficients of the linear regression model for the selected brands (Topshop and Dior). The coefficients indicate the change in the dependent variable (log of number of likes) for a one-unit increase in the independent variables (number of new celebrities and number of traditional celebrities).

### For Topshop, the coefficient of the number of new celebrities is 0.0023, which means that for every one-unit increase in the number of new celebrities, the log of the number of likes increases by 0.0023. The coefficient of the number of traditional celebrities is 0.039, which means that for every one-unit increase in the number of traditional celebrities, the log of the number of likes increases by 0.039.

### For Dior, the coefficient of the number of new celebrities is 0.054, which means that for every one-unit increase in the number of new celebrities, the log of the number of likes increases by 0.054. The coefficient of the number of traditional celebrities is 0.053, which means that for every one-unit increase in the number of traditional celebrities, the log of the number of likes increases by 0.053.

### Dior:

### The output of the linear regression for Dior shows that the coefficients for the number of new celebrities and traditional celebrities are -0.01011589 and 0.04058056, respectively. This means that for every one unit increase in the number of new celebrities, the natural logarithm of the number of likes received decreases by 0.01011589 units, and for every one unit increase in the number of traditional celebrities, the natural logarithm of the number of likes received increases by 0.04058056 units

1. [↑](#footnote-ref-1)