**User Engagement Activity Analysis**

**The Company:**

*Zen* is a business-to-business professional services company that provides online services and products to support work of the employees of its customer companies. Many companies in diverse areas of business including, for example, manufacturers, producers, suppliers, retailers, transportation, and others are *Zen’s* paid subscribers. Besides an annual subscription fee, variable charges are based on the extent of engagement with its products and services by the users (users are employees of customer companies).

*Zen* has a Data Analytics (DA) department, whose primary responsibility is to track user engagement and support better product and business decisions using data. The DA teams conduct studies, carry out projects to address specific business problems, and perform ad-hoc analyses to support business decisions. You are one of the business analysts in the DA department.

*Zen* defines user activity as an engagement with its online portal, i.e., the customers (users) having made some type of server call by interacting with the company’s website/web server. Such events are listed as “engagement” in the event\_type column of the EVENTS table.

**Task at Hand:**

Your job is to monitor user activity and engagement on a regular basis and report to the senior management, whether engagement is stable, dropping, or increasing. For any changes in user engagement activity, you also need to investigate possible sources or reasons for the changes.

**Data:**

Data is provided in three csv files, which you’ll need to upload to your Databricks account and create permanent tables so you can use SQL with those tables. The structure of the three csv files is described below.

|  |  |
| --- | --- |
| USERS | |
| user\_id: | A unique ID per user. Can be joined to user\_id in either of the other tables. |
| created\_at: | The time the user was created (first signed up) |
| state: | The state of the user (active or pending) |
| activated\_at: | The time the user was activated, if they are active |
| company\_id: | The ID of the user's company |
| language: | The chosen language of the user |

|  |  |
| --- | --- |
| EVENTS | |
| user\_id: | The ID of the user logging the event. Can be joined to user\_id in either of the other tables. |
| occurred\_at: | The time the event occurred. |
| event\_type: | The general event type. There are two values in this dataset:  signup\_flow: refers to anything occurring during the process of a user's authentication,  engagement: refers to general product usage after the user has signed up for the first time. |
| event\_name: | The specific action the user took - possible values include the following:  create\_user: User is added to the company’s database during signup process  enter\_email: User begins the signup process by entering her email address  enter\_info: User enters her name and personal information during signup process  complete\_signup: User completes the entire signup/authentication process  home\_page: User loads the home page  like\_message: User likes another user's message  login: User logs into her/his account  search\_autocomplete: User selects a search result from the autocomplete list  search\_run: User runs a search query and is taken to the search results page  search\_click\_result\_X: User clicks search result X on the results page, where X is a number from 1 through 10.  send\_message: User posts a message  view\_inbox: User views messages in her inbox |
| location: | The country from which the event was logged (collected through IP address). |
| device: | The type of device used to log the event. |

|  |  |
| --- | --- |
| EMAILS | |
| user\_id: | The ID of the user to whom the event relates. Can be joined to user\_id in either of the other tables. |
| occurred\_at: | The time the event occurred. |
| action: | The name of the event that occurred.  sent\_weekly\_digest: it means that the user was delivered a digest email showing relevant conversations from the previous day.  email\_open: it means that the user opened the email.  email\_click-through: it means that the user clicked a link in the email. |