# **Contoh Soal**

**Animal Crossing Achievement**

**Animal Crossing** is a popular simulation genre pocket game created by Nintendo. **ACNL** is a Fan Based Club for Animal Crossing. This time, ANCL want to make a data mining about Animal Crossing’s achievements. Some of the achievement need to be done in sequence. There are several tiers of reward for each achievement, and you can get a specific reward based on Tier that you reach. This ANCL has data for Animal Crossing which are provided in **csv** format as following:

* **achievements-category**.**csv**

|  |  |  |
| --- | --- | --- |
| **Attribute** | **Data Type** | **Description** |
| Unique Entry ID | Character | Unique ID in the Achievement. |
| Category | Character | Recorded Achievement’s Category in the game |

* **achievements**.**csv**

|  |  |  |
| --- | --- | --- |
| **Attribute** | **Data Type** | **Description** |
| Name | Character | Achievement’s name in game. |
| Award Criteria | Character | Criteria to get the Achievement. |
| Num of Tiers | Integer | Total Tier of Achievement. |
| Tier 1 | Integer | Required Point to Tier 1. |
| Tier 2 | Integer | Required Point to Tier 2. |
| Tier 3 | Integer | Required Point to Tier 3. |
| Tier 4 | Integer | Required Point to Tier 4. |
| Tier 5 | Integer | Required Point to Tier 5. |
| Reward Tier 1 | Integer | Reward for Tier 1. |
| Reward Tier 2 | Integer | Reward for Tier 2. |
| Reward Tier 3 | Integer | Reward for Tier 3. |
| Reward Tier 4 | Integer | Reward for Tier 4. |
| Reward Tier 5 | Integer | Reward for Tier 5. |
| Reward Tier 6 | Integer | Reward for Tier 6. |
| Sequential | Character | Is achievement sequential. |
| Version | Character | Achievement release version. |
| Unique Entry ID | Character | Unique ID in the Achievement. |

As the data scientist in ANCL, you are asked to do some **tasks** below:

1. **Data Visualization**

To help the player understand the data easier, you are asked to **visualize** the data in **graph** **form**. Some data that is needed to be visualized are:

* 1. Show the **top 5** most frequent **achievement** based on **category**.

Chart, pie chart

Description automatically generated

**Figure 1. Top 5 Achievement Category Chart**

* 1. Show the **frequency** of **achievement** basedon **sequential**.

Chart, bar chart

Description automatically generated

**Figure 2. Frequency of Sequential Achievement Chart**

* 1. Show the **total number of achievement** basedon **reward tier 1.** Separate the achievement into **3 categories** as below:
  + **Large**, where reward tier 1 is **greater than 820**.
  + **Normal**, where reward tier 1 is **greater than 480** and **less than or equal to 820**.
  + **Small**, where reward tier 1 is **less than or equal to 480**.

Chart, bar chart

Description automatically generated

**Figure 3**. **Total Number of Achievement based on Reward Tier 1 Chart**

1. **Frequent Pattern Analysis**

To help players know what **achievement category** are frequently appeared in **each** **tier**, you are asked to do **frequent pattern analysis** to search for **frequent achievement categories** within a tier. To get the frequent achievement category data, use the **achievements.csv** and **achievements-category**.**csv** file.

* 1. **Data pre-processing**

In the data pre-processing phase, there are some data that can’t be used for the further analysis. Do the following task to **cleanse** the data:

* **Remove** all achievement which **category** is ‘**Money**’.
* **Remove** all achievement which **category** is ‘**Communication**’**.**
* **Remove** all achievement which **category** is ‘**LandMaking**’.
* **Remove** allachievement which **sequential** is ‘**No**’.
* **Remove** allachievement which **tier** is ‘**6**’.
  1. **Data transformation**

In this phase, you need to **change** **the** **data**, so it is suitable to be used in the **apriori** **analysis**. Prepare the data in terms of the **achievement’s category**.

* 1. **Data mining**
* Show **frequent** **achievement’s** **category** using **apriori** algorithm with **minimum support** of **0**.**6** based on the data that have already pre-processed.

Table

Description automatically generated

***Figure 4. Frequent Achievement’s Category using Apriori***

* Show the **association** **rules** using **minimum confidence** of **0**.**8** based on the frequent **achievement’s category** that resulted from step above.

Table

Description automatically generated

***Figure 5. Association Rules***

**References**

* <https://www.kaggle.com/jessicali9530/animal-crossing-new-horizons-nookplaza-dataset?select=achievements.csv>