Tables in RatingCrawler

Table I: AppData

This table is used to store the overall information about an app.

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Type | Key | Description |
| Id | int(11) | Y | Auto increment |
| Appid | Varchar(100) |  | Packet name in google play |
| Developer\_name | Varchar(100) |  |  |
| Developer\_id | Varchar(50) |  |  |
| Developer\_comany | Mediumblob |  | The company of the developer. |
| Average\_rating | Float |  |  |
| Total\_raters | Int(11) |  |  |
| Rated\_summary | Varchar(50) |  | The summary of the ratings like   |  | | --- | | Frequent/Intense Cartoon or Fantasy Violence; | |
| 1star\_num | Int(11) |  |  |
| 2star\_num | Int(11) |  |  |
| 3star\_num | Int(11) |  |  |
| 4star\_num | Int(11) |  |  |
| 5star\_num | Int(11) |  |  |
| Current\_version | Varchar(50) |  | The latest version of the current app. |
| Required\_android\_version | Varchar(100) |  | The minimum required Android version by the app. |
| Category | Varchar(50) |  | The category of current app.  Like games |
| Sub\_category | Varchar(50) |  | Like racing in games |
| Price | Float |  | The price of this app |
| Currency | Varchar(50) |  | The currency of the price. |
| Rank | Int(11) |  | This rank is actually different each day. |
| Thumbnail | Varchar(200) |  | The link to the app’s icon. |
| Store\_location | Varchar(100) |  | The location of current store  For apple, it may be China like 14345. See the appendix. |
| App\_name | mediumblob |  | The name of the app name. It may contain some strange symbols. |
| Crawler\_point | Varchar(100) |  | The entry point of the crawler like top-paid, top free of each category |
| Description | Text |  | To be tested. |
| Latest\_modified | Datetime |  | The time of the latest update. |
| Latest\_action | Varchar(50) |  | The latest action of the developer such as “update” or “released”. |
| Installs | Int(11) |  | The installations of this app. |
| Content\_rated | Varchar(20) |  | 4+ like in apple  Everyone or low maturity in google play |
| Language | Text |  | The language this app could support. |
| Requirements | mediumblob |  | The requirement of this app.   |  | | --- | | “Compatible with iPhone 3GS, iPhone 4, iPhone 4S, iPhone 5, iPod touch (3rd generation), iPod touch (4th generation), iPod touch (5th generation) and iPad. <BR/>Requires iOS 4.0 or later. “ | |
| url | Mediumblob |  | The url of this application. It might be null if the app has been removed. |
| Latest\_review\_id | Varchar(20) |  | The lasted reviewer’s id. It is for comment update. |
| Size | Varchar(20) |  | The size of the app. |
| Permission | text |  |  |
| Contact\_phone | text |  | The contact of the author. |
| Contact\_email | text |  | The contact email address of the developer. |
| Contact\_website | Text |  | The contact website of the developer. |
| Promoted\_text | Text |  | The first paragraph in description |
| Promoted\_screenshot | Text |  | The images addresses of app screenshot. |
| Promoted\_video | Text |  | The link to the video promotion. |
| Recent\_changes | Text |  | Like what’s new in google play |
| Note | Text |  | For what? |
| *Rho* | *Float* |  | *The correlation coefficient.* |
| *Adjusted\_rating* | *Float* |  | *The estimated app’s quality.* |
| *Average\_weekly* | *Float* |  | *The average rating of weekly average rating.* |
| *Week* | *Int(11)* |  | *The life time of the app with the number of weeks.* |

Note:

1. The column with blue background is only for special purposes.

Table II: Comment

This table is for storing comment information.

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Type | Key | Description |
| Id | int(11) | Y | Auto increment |
| Appdata\_id | Int(11) |  | The id in the table AppData. |
| Reviewer | Mediumblob |  | The name of the reviewers. |
| Reviewer\_id | Varchar(100) |  | The id of the reviewer. |
| Date | Datetime |  | The releasing date of this comment. |
| Device | Varchar(100) |  | The type of the device(only apply to google).  1: Samsung Nexus S with version 2.0.1 |
| Device\_version | Mediablob |  | The version the app in the reviewer’s device. |
| Rating | Int(11) |  | The rating like 5 star. |
| Comment | Mediumblob |  | The text comment. |
| Comment\_title | Mediumblob |  | The text comment title. |
| Helpfulness\_ratio | Float |  | The percentage of agreement. |
| Helpfulness\_agree | Int(11) |  | The number of reviewers agree with this comment. |
| Helpfulness\_total | Int(11) |  | The total number of reviewers which express their agreement. |
| Review\_id | Int(11) |  | The id of this review. |
| App\_id | Int(11) |  | The app id which this review aimed at. |
| *Contribution* | *Float* |  | *The contribution of this rating* |

Table III: AppRatings

This table is used to store all of the rating in one month. At each time, you can check the ranking information in previously month.

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Type | Key | Description |
| Id | int(11) | Y | Auto increment |
| App\_id | Int(11) |  | The id of the app. |
| App\_name(title) | Mediumblob |  | The name of the app. |
| Update\_time | Datatime |  | The update time of the latest update. |
| Day1…31 | Int(11) |  | The rank in the first day of the month.  To the last day of the month. |

Table IV: Reviewers

For storing all the apps reviewed by one reviewer.

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Type | Key | Description |
| Id | int(11) | Y | Auto increment |
| Reviewer\_id | Varchar(50) |  | The id for this reviewer. |
| App\_ids\_ordered | Text |  | The apps’ id which this reviewer have left a comment. All of the apps’ id are ordered in increment. |
| Review\_ratings | Text |  | The rating of each review. |
| Review\_dates |  |  | The date of releasing each review. |
| Size |  |  | How many apps the reviewer has reviewed. |
| Developers | Text |  | The developer’s name |
| Cluster\_id |  |  | The cluster this reviewer belongs to. |
| Status |  |  | The status of this reviewer. |

Table V: Developers

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Type | Key | Description |
| Id | int(11) | Y | Auto increment |
| Developer\_id | Varchar(50) |  | The id of the developer. |
| App\_id | Varchar(50) |  |  |