Here's the final, compiled schema for your channels table in Firestore, including the categories for llm\_thumbnail\_skin\_color.

**Firestore channels Table Schema**

Each document in your channels collection (e.g., /artifacts/{appId}/public/data/channels/{channelId}) will follow this structure:

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Data Type** | **Description** | **Source / Notes** |
| channel\_id | string | Unique identifier for the YouTube channel. **Primary Key.** | YouTube Data API (id) |
| channel\_name | string | The public display name of the YouTube channel. | YouTube Data API (snippet.title) |
| channel\_url | string | Direct URL to the YouTube channel page. | Derived (e.g., https://www.youtube.com/channel/{channel\_id}) |
| channel\_description | string | The channel's "About" section description. | YouTube Data API (snippet.description) |
| published\_at | timestamp | Date and time the channel was created. | YouTube Data API (snippet.publishedAt) |
| profile\_picture\_url | string | URL to the channel's profile image (e.g., high resolution thumbnail). | YouTube Data API (snippet.thumbnails.high.url) |
| subscriber\_count | number | Total number of subscribers (at time of data collection). | YouTube Data API (statistics.subscriberCount) |
| video\_count | number | Total number of public videos on the channel. | YouTube Data API (statistics.videoCount) |
| total\_view\_count | number | Total views across all videos on the channel. | YouTube Data API (statistics.viewCount) |
| country | string | The country associated with the channel, if specified by the owner. | YouTube Data API (snippet.country) (Optional: May not always be present) |
| topic\_ids | array<string> | List of Freebase Topic IDs broadly associated with the channel's content (e.g., /m/041xxh for beauty). | YouTube Data API (topicDetails.topicIds) |
| last\_content\_upload\_at | timestamp | Timestamp of the most recent video upload to the channel. Critical for filtering active channels. | **Derived:** Fetch playlistItems from the channel's uploads playlist, sorted by publishedAt in descending order, then take the timestamp of the first item. |
| last\_updated | timestamp | Timestamp of when this channel's *metadata in your database* was last fetched/processed from the YouTube API. | Your application (for freshness checks of the channel record itself) |
| features\_aligned\_with | array<string> | **Annotated/LLM-Inferred (from video/content analysis):** Facial features of the *influencer* or features their tutorials primarily cater to. | Derived from LLM analysis of channel description, video titles/descriptions, and video transcripts. **(See Predefined Categories below)** |
| primary\_makeup\_style | array<string> | **Annotated/LLM-Inferred (from video/content analysis):** Dominant makeup style(s) they showcase. | Derived from LLM analysis of channel description, video titles/descriptions, and video transcripts. **(See Predefined Categories below)** |
| target\_skill\_level | array<string> | **Annotated/LLM-Inferred (from video/content analysis):** Skill levels their tutorials primarily target. | Derived from LLM analysis of channel description, video titles/descriptions, and video transcripts. **(See Predefined Categories below)** |
| demographic\_focus | array<string> | **Annotated/LLM-Inferred (from video/content analysis):** Specific demographics they cater to (optional). | Derived from LLM analysis of channel description, video titles/descriptions, and video transcripts. **(See Predefined Categories below)** |
| llm\_thumbnail\_face\_shape | string | **LLM-Inferred (from thumbnail image):** Face shape detected in the channel's thumbnail. | Gemini AI multimodal analysis (from profile\_picture\_url). Value will be one of the face\_shape categories or unknown. |
| llm\_thumbnail\_skin\_color | string | **LLM-Inferred (from thumbnail image):** Skin color detected in the channel's thumbnail. | Gemini AI multimodal analysis (from profile\_picture\_url). Value will be one of the skin\_color categories or unknown. |
| llm\_thumbnail\_eye\_shape | string | **LLM-Inferred (from thumbnail image):** Eye shape detected in the channel's thumbnail. | Gemini AI multimodal analysis (from profile\_picture\_url). Value will be one of the eye\_shape categories or unknown. |
| llm\_thumbnail\_lip\_shape | string | **LLM-Inferred (from thumbnail image):** Lip shape detected in the channel's thumbnail. | Gemini AI multimodal analysis (from profile\_picture\_url). Value will be one of the lip\_shape categories or unknown. |
| llm\_thumbnail\_status | string | Status of the thumbnail image analysis process. | Your application's LLM processing logic. **(See Predefined Categories below)** |
| llm\_thumbnail\_analysis\_date | timestamp | Timestamp of when the thumbnail analysis was last performed. | Your application's LLM processing logic. |

**Predefined Categories for Specific Columns:**

1. features\_aligned\_with (array&lt;string>)

(This should be a combination of categories from the shapes/types below that the channel focuses on or that the influencer visibly possesses and features in their content)

* **Face Shape:**
  + round\_face
  + oval\_face
  + square\_face
  + heart\_face
  + long\_face
  + diamond\_face
* **Eye Shape:**
  + hooded\_eyes
  + almond\_eyes
  + round\_eyes
  + downturned\_eyes
  + monolid\_eyes
  + up-turned\_eyes
  + deep\_set\_eyes
  + close\_set\_eyes
  + wide\_set\_eyes
* **Skin Undertone:**
  + warm\_undertone
  + cool\_undertone
  + neutral\_undertone
* **Skin Type/Concerns:**
  + oily\_skin
  + dry\_skin
  + combination\_skin
  + acne\_prone\_skin
  + sensitive\_skin
  + mature\_skin
* **Lip Shape:**
  + fuller\_lips
  + thin\_lips
  + heart\_shaped\_lips
  + bow\_shaped\_lips
* **Other:**
  + narrow\_nose
  + wide\_nose
  + high\_cheeks
  + low\_cheeks

**2. primary\_makeup\_style (array&lt;string>)**

* natural
* everyday
* glam
* editorial
* bridal
* dramatic
* minimalist
* soft\_glam
* no\_makeup\_makeup

**3. target\_skill\_level (array&lt;string>)**

* beginner
* intermediate
* advanced
* pro\_artist

4. demographic\_focus (array&lt;string>)

(Optional, use only if channels explicitly target these)

* teen\_focused
* adult\_focused
* mature\_age\_focused
* dark\_skin\_tones
* light\_skin\_tones
* male\_makeup
* non\_binary\_makeup

**5. llm\_thumbnail\_face\_shape (string)**

* Round
* Oval
* Square
* Heart
* Long
* Diamond
* unknown

**6. llm\_thumbnail\_skin\_color (string)**

* fair
* light
* medium
* tan
* dark
* deep
* unknown

**7. llm\_thumbnail\_eye\_shape (string)**

* Almond
* Round
* Hooded
* Monolid
* Downturned
* Up-turned
* Deep-set
* Close-set
* Wide-set
* unknown

**8. llm\_thumbnail\_lip\_shape (string)**

* Fuller
* Thin
* Heart
* Bow
* unknown

**9. llm\_thumbnail\_status (string)**

* analyzed
* skipped\_multiple\_faces
* skipped\_no\_clear\_face
* error\_processing\_image
* error\_llm\_response
* pending
* unknown