

Exploratory Data Analysis Report of Coursera, Udemy, and Youtube

Rye International Data Team presents for Onomy

1. Key Takeaways

- a. **Video length:** Onomy's video lengths are shorter compared to Udemy and YouTube; Advertisements are relatively longer for current courses' lengths.
Remark: Nevertheless, TikTok might be a good platform to post short videos
- b. **Course Topic:**
 - i. Three platforms have relatively few intersections of Onomy's intended adulting business, leaving a potentially large market.
 - ii. Across the three major online platforms, there are fewer courses in categories such as Healthcare, mortgage, tax.
 - iii. Personal-finance-related courses are dominant among competitors.
- c. **Course Content:** For popular education Youtuber, colorful cartoons are often used throughout these course videos besides the view of the lecturer. The volume of those videos are high compared to Onomy's videos with bullet points listed aside.

2. Brief summary of the whole report

- a. Competitor analysis on Coursera, which includes numbers of each relevant popular courses segments, ratings of different related segments, number of subscribers for popular segments, relationship/overlap between different segments.
- b. Competitor analysis on Udemy, including the numbers of course categories, number of subscribers comparison, the average number of subscribers gained per course, comparison, and trends in time length.
- c. Competitor analysis on Youtube, including the category composition of all education channels, the identification of popular education channels and course-related channels to find their commons in course structure and length.

3. Motivation / Purpose of conducting Exploratory Data Analysis

- a. **Motivation:** We try to understand the popular trends of online education platforms, including their popular coursework, ratings, etc. We hope to gain a better understanding of market demand, and users' preferences through analyzing Onomy's potential competitors and provide insights for Onomy's potential course designs, such as topics and video length.
- b. **Questions we try to answer:**
 - i. Among all courses, what proportion is about Onomy's current topics?
 - ii. What type of courses are more popular than others?
 - iii. What are the intersections between different categories of courses?
 - iv. What type of courses have the longest/shortest length of the videos?

4. *How do we conduct EDA? Where is the data from? analysis methods?*

a. Data source:

- i. **Coursera:** we obtain two datasets from Kaggle. One data set contains nearly all courses from Coursera, with course name, ratings, difficulty level, tags, course URL. Another dataset contains a smaller set of courses from Coursera, with course title, course organization, course-Certificate-type, course-rating, course-difficulty, course_students_enrolled.
- ii. **Udemy:** We obtain two datasets from Data.world and Kaggle. One data set contains all courses relevant to Finance and Accounting from Udemy as of September 2020, containing course name, ratings, number of subscribers. The other is a smaller data set that focuses on Business and Finance courses, containing course name, number of subscribers, difficulty, number of lectures, and time length.
- iii. **Youtube:** From Kaggle, we get ‘Top 100,000 Youtube Channels’ based on number of followers, including their category, title, description, number of followers and number of videos. The dataset is published in 2019.

b. Analysis process:

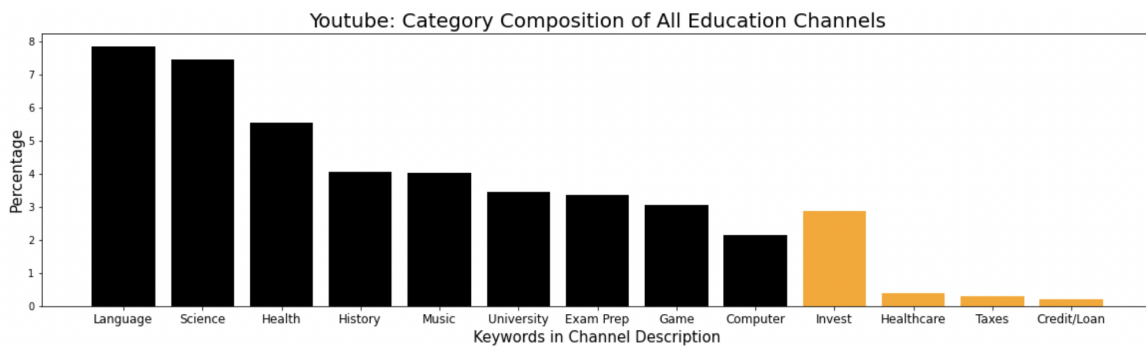
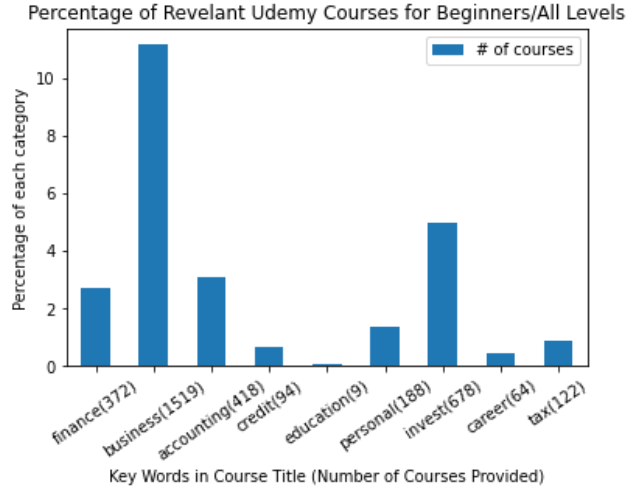
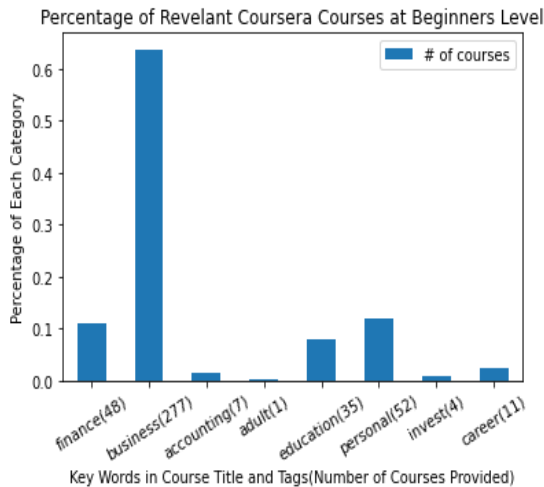
- i. For the Coursera dataset, we filter out courses for non-beginner level, non-English courses, courses that have no ratings, or courses that have no valid course enrollments numbers. Then we divide the desired dataset into different categories by searching for keywords in the course name, course tags, and course organization. Next we analyze the needed parameters based on different categories’ datasets.
- ii. A similar procedure is performed for the data sets for Udemy. We filter out the courses that require prior experiences (the intermediate and expert level courses) and divide the dataset into different categories by looking for keywords in the title.
- iii. For the Youtube dataset, we filter out all the Education Channels and remove those related to kindergarten education. After dropping duplicates, non-english channels, and those with empty description, we end up with 4441 channels.

5. *What are the results and implications?*

a. The Number of Courses provided in different categories.

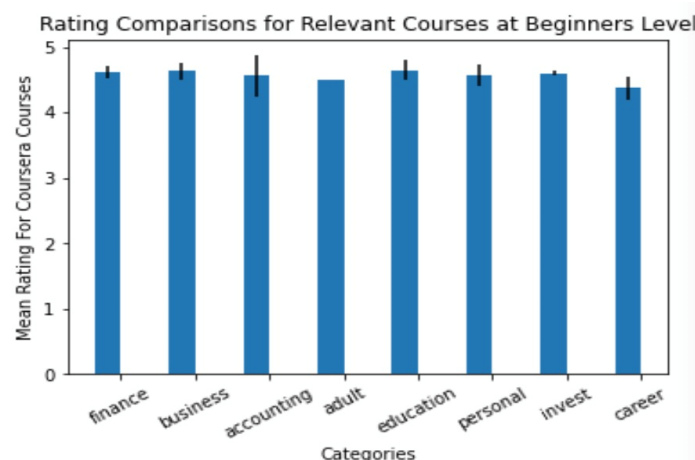
- i. We can see that business-related courses are dominant among all categories, indicating the number of courses on business-related is already high. This might suggest that if Onomy designs courses on how to develop its own business for adults, it may face fierce competition.
- ii. We can see that very few courses directly contain Adults, which might imply that people do not have the demand to watch adulting content provided by academy-focused platform Coursera.

- iii. On Coursera, no/very few entry levels courses for personals at these segments: tax, credit cards, insurance.
- iv. On Udemy, there are relatively few courses for the categories of personal, education, tax, credit, and career
- v. For Youtube, the education channel is dominated by academic courses in science and language learning. Among Onomy's topics, investing is the most popular one and there are very few related to tax, credit cards, and insurance.



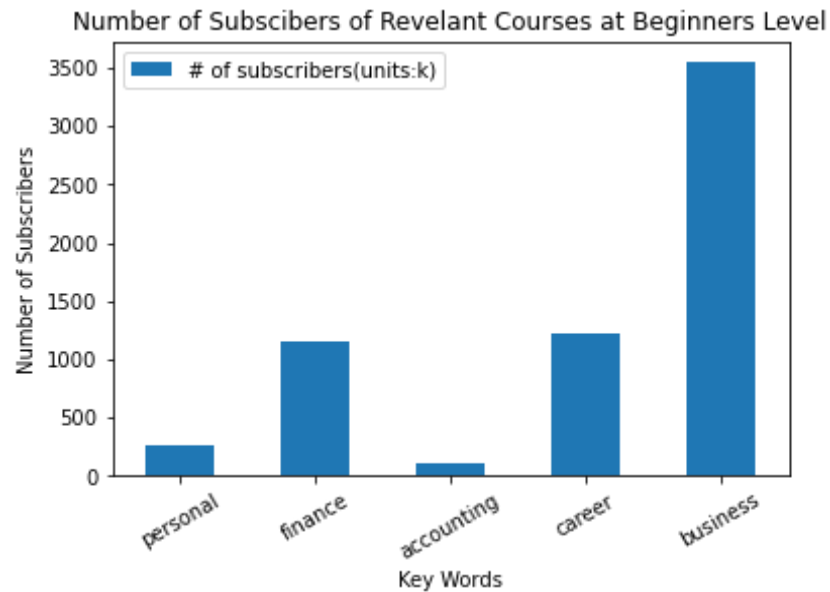
b. Showcase of popular courses (based on # of subscribers/ratings)

- i. **Rating comparison:** Courses provided by Coursera overall have relatively similar ratings. This poses an opportunity or hopes to expand the market in online adulting education for Onomy.

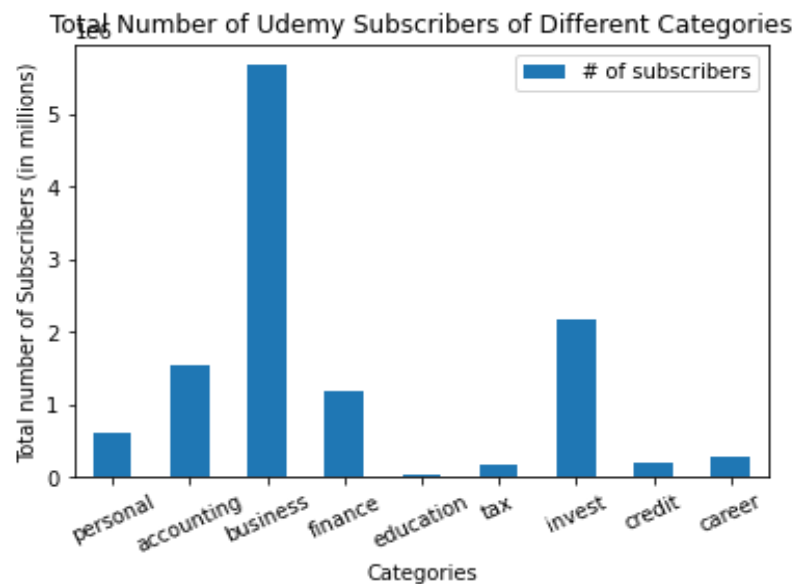


- ii. **Number of Subscribers:** For Coursera, business related courses have the most subscribers in total due to the high number of coursework provided and perhaps high demand. On the other hand, personal content shows a relatively low number of subscribers, which might pose some hope to online personal education but might also indicate low demand.

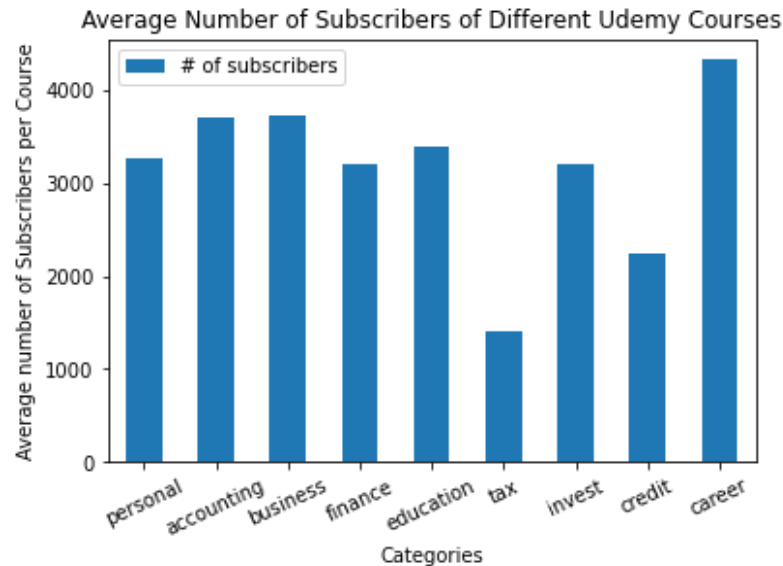
For Coursera (data obtained from the second datasets which does not contain courses that have keywords of personal/adult):



For Udeemy:



- iii. **The average number of subscribers:** Interestingly, from our data in Udemy, the career category has the highest average number of subscribers per course, whereas the tax and credit categories have the lowest average number of subscribers.



- iv. **Analysis of popular courses in Youtube:** Most popular courses-channels have video length about 5 to 10 minutes which might be both informative and attractive for viewers; Colorful cartoons are often used throughout these course videos besides the view of the lecturer; The volume of the video is large with bullet points listed aside; Very few Youtubers focused on taxes, healthcare, and credit cards.

a) **Invest**

Popular Youtubers: Tai Lopez, Ricky Gutierrez, Grant Cardone

Features: millionaires sharing their personal investing ideas at home/office/personal plane. People are attracted by their success

Remarks: Tai Lopez is selling his courses “67 LESSONS THAT SCHOOLS FAILED TO TEACH US TO FOLLOW IN 2021.”

b) **Healthcare**

Popular Youtubers: Healthcare Triage

Features: Dr. Aaron Carroll explains healthcare policy, medical research, medicine, health, and healthcare. Each video is about 5 to 8 minutes with colorful cartoons and charts.

c) **Taxes/Credit/Loan**

Popular Youtubers: MoneyCoach, ClearValue Tax

Features:

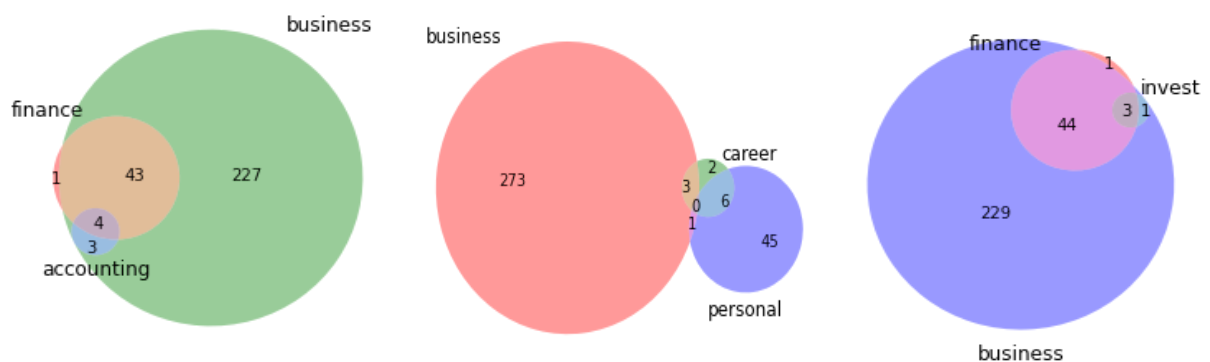
- 1) MoneyCoach covers almost all of Onomy’s topics, including taxes, investing, credit, and healthcare. Each video is about 2 to 4 minutes and the most popular one receives about 510,000 views. All content is conveyed by drawing.

- 2) ClearValue Tax is a channel that has gotten really popular in the past two years. It helps people to do tax preparation and share stimulus payment news. People have high demand for this, but very few Youtubers focus on it.

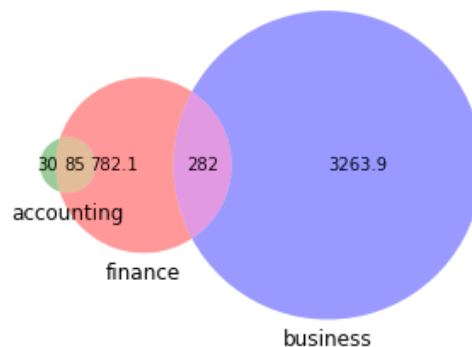
c. Relationship between categories:

From both data sets, we can see the number of subscribers or courses provided have a larger intersection between business and other finance-related coursework. The personal category has a slightly few intersections with career-related coursework. Onomy may gain product insights such as how different themes can be incorporated into another and how to develop a series of courses that provide diverse contents but share similar themes.

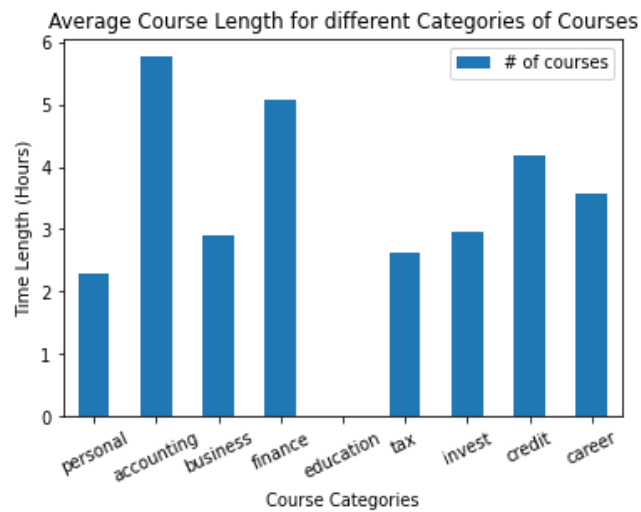
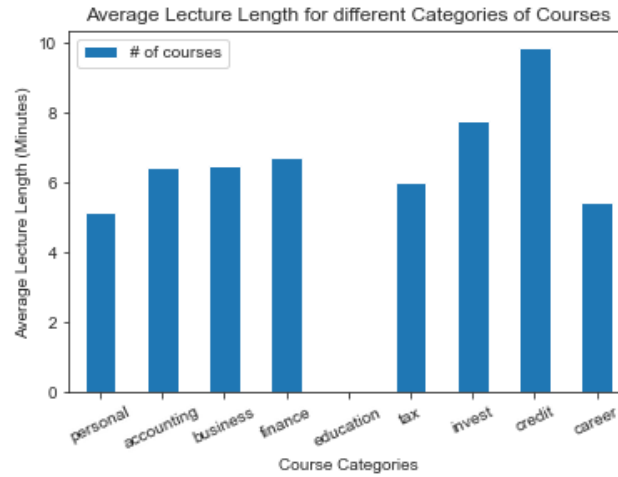
i. Number of courses in different categories and their intersection:



ii. Number of subscribers in different categories and their intersection:



- iii. **Time length of different categories:** From UdeMy's second data set, we found that Accounting and Finance tend to have the longest courses; Personal related courses tend to have the shortest courses. Credit related courses tend to have the longest lecture; Personal related courses tend to have the shortest lectures. (Average Lecture Length computed from the mean of the course length divided by number of published lectures)



Appendix

1. A few Potential Relevant Courses For Onomy (for non-professionals and personal)

Obtained through filtering and directly looking at Coursera course names. Onomy can potentially gain some product insights by looking at their contents design, course structure, and other product features.

a. business-related:

- Financial Accounting Fundamentals
- **Accounting for Decision Making**
- Accounting in 60 Minutes - A Brief Introduction

b. career:

- **Improving Communication Skills**
- **Blended Learning: Personalizing Education for Students**
- **How to Write a Resume (Project-Centered Course)**
- **Creative Thinking: Techniques and Tools for Success**
- Career Brand Development and Self-Coaching
- Success
- Speaking to inform: Discussing complex ideas with clear explanations and dynamic slides
- IBM IT Assessment: Identifying the Right Career for You!
- Essentials for English Speeches and Presentations
- Personality Types at Work
- Effective Communication in the Globalised Workplace - The Capstone

c. personal life:

- **Financial Planning for Young Adults**
- **International Travel Preparation, Safety, & Wellness**
- Empowering Yourself in a Post-Truth World
- What does it mean to identify as Transgender or Gender Non-Conforming (TGNC)?
- The Arts and Science of Relationships: Understanding Human Needs
- Healing with the Arts

2. Overview of Popular Education Channels in Youtube

a. Top 5 Education Channels:

We identify the 5 most popular education channels based on their number of followers. They are 'The King of Random', 'CrashCourse', 'TED-Ed', 'Matthew Santoro', and 'Planet Dolan'.

Below are our short summary of their content:

The King of Random (aka TKOR)

a place where curiosity, creativity, and experimentation meet. We're all about learning how things work, doing cool projects, and sharing our discoveries with you. It joined Youtube around 2010, uploading 20 videos per month, each about 10 min.

Crash Course

produced more than 32 courses on a wide variety of subjects, including organic chemistry, literature, world history, biology, philosophy, theater, ecology, and many more! It joined Youtube around 2006, uploading 7 videos per month. **Including many courses for investing/taxes/insurance, each about 10 min long. Might be Onomy's competitor. But most of the content is academic-oriented. For example, when it talks about taxes, the main content is about different tax systems instead of how to fill in the tax form. The volume of the video is large with colorful cartoons and change in voices/views.**

TED-Ed

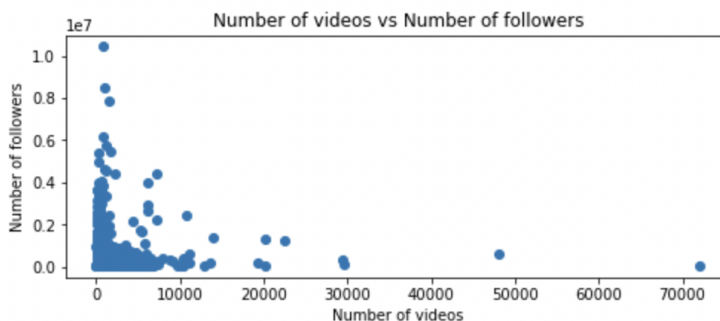
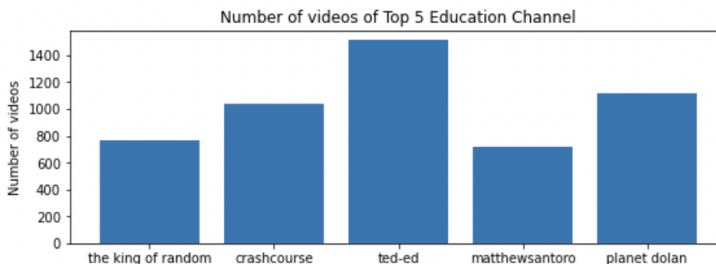
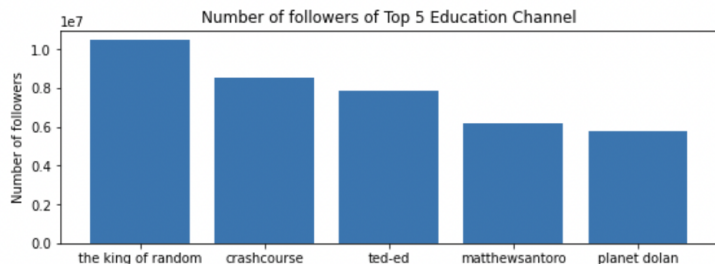
creating lessons worth sharing is an extension of TED's mission of spreading great ideas. It joined Youtube about 2010, with 10 videos per month. **The channel covers topics about investing, but not the rest. Each video is about 5 min with cartoon-like pictures.**

Matthew Santoro

a Canadian YouTuber and comedian. He produces top ten lists and "50 Amazing Facts" videos on his main channel. It joined Youtube about 2010, with 3 to 6 videos per month, each about 10-20 min.

Planet Dolan

answers fan-submitted questions in a way that is completely inaccurate (most of the time), but is meant to convince the audience of its validity, or simply for entertainment purpose. It joined Youtube about 2014, with 6 videos per month, each about 9 min.



Implications:

1. The number of videos ranges from 700 to 1500 for these 5 channels. Though there is no explicit relationship between the number of followers and number of videos, these top

channels update relatively fast with more than 6 videos per month. Also, most of their videos are from 5 min to 10 min to keep viewers from distraction.

2. Among the 5, CrashCourse is the one that is closest to Onomy's topic. We might want to learn more about their strategies of attracting viewers and their structure of courses.

b.Top 5 Course-related Channels:

We identify the 5 most popular course-related channels based on their number of followers. Besides, 'CrashCourse' and 'TED-Ed', we find 'Khan academy', 'Learn English with Let's Talk - Free English Lessons', and 'Unacademy'.

Below are our short summary of their content:

Khan Academy

Our interactive practice problems, articles, and videos help students succeed in math, biology, chemistry, physics, history, economics, finance, grammar, and many other topics. It joined Youtube in 2006, uploading 20 videos per month, each with 5min. **Including many courses about investing/taxes/insurance, each about 10 min long. Might be Onomy's competitor.**

Learn English with Let's Talk - Free English Lessons

India platform of learning English for Business, IELTS, and TOFEL. Uploading 10 videos per month, each about 10 min. Their topics have almost no overlap with Onomy.

Unacademy

India's largest free education initiative focuses on academic teaching. Their topics have almost no overlap with Onomy.

Implication:

1. Based on content summaries above, we identify 'CrashCourse' and 'Khan Academy' as two potential competitors of Onomy. It could be insightful to investigate their popular courses about investing, healthcare, credit cards and taxes to find their traits and identify what makes viewers like them. Onomy could then tailor its strategies based on these findings.
2. Also, both platforms provide videos with about 10 min long. Maybe Onomy could consider extending their video length