

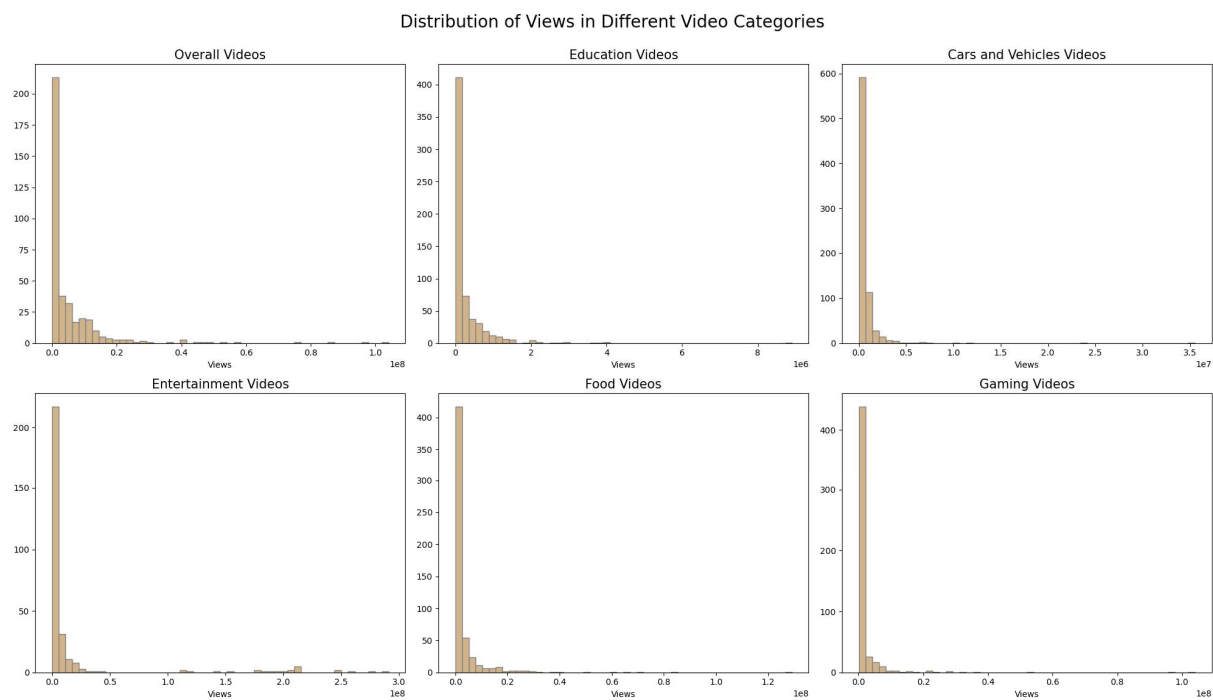
# Supplementary Information to “How Video Length Influences the Number of Viewers and their Engagement”

Hanna Sophie Müller

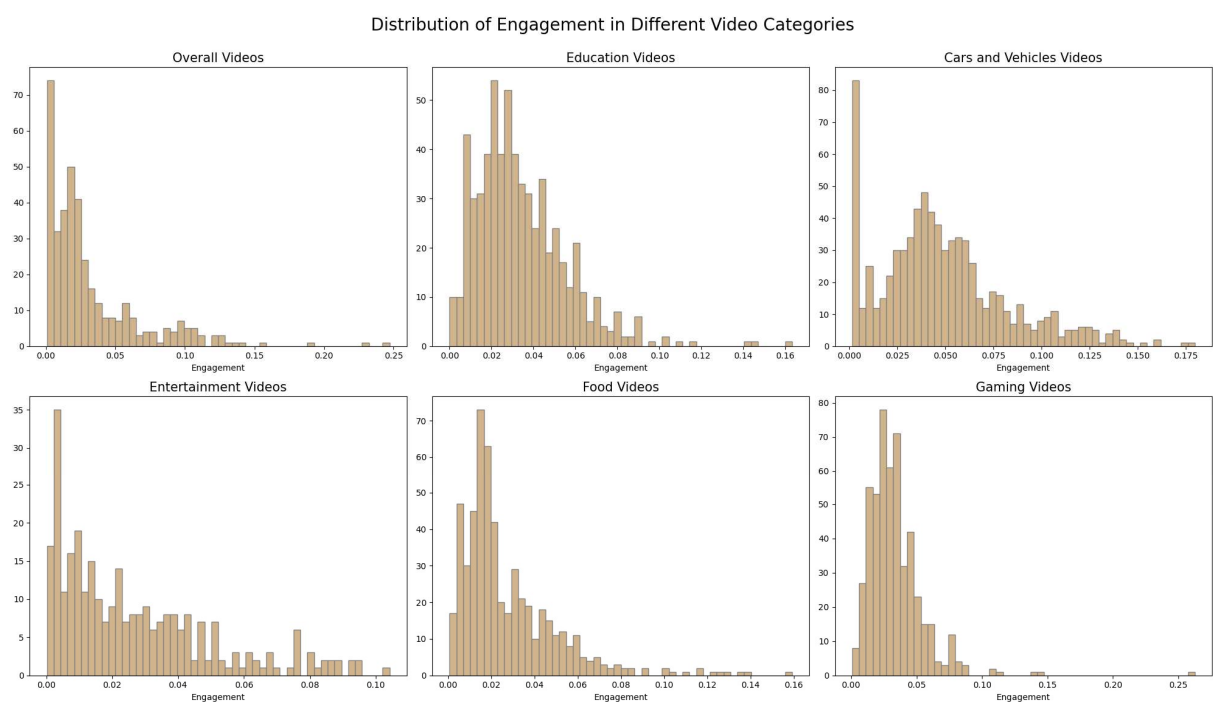
University of Konstanz

Matriculation Number: 01/1358625

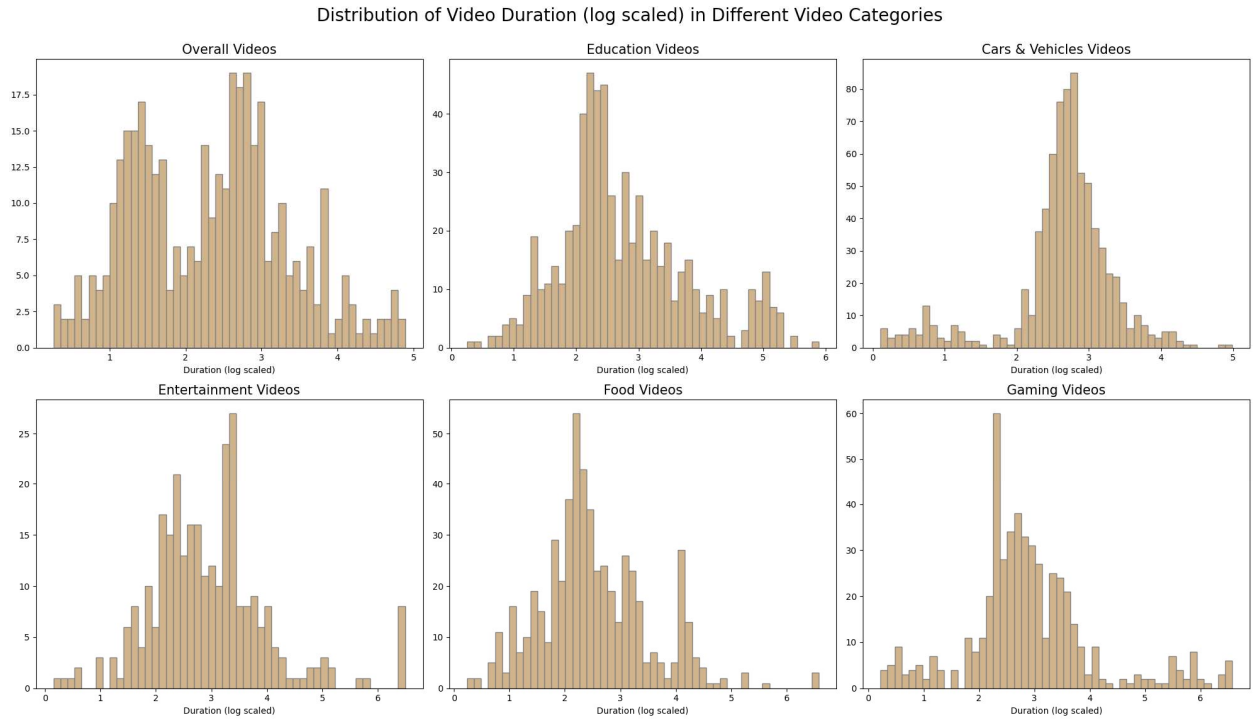
E-Mail: [hanna.3.mueller@uni-konstanz.de](mailto:hanna.3.mueller@uni-konstanz.de)



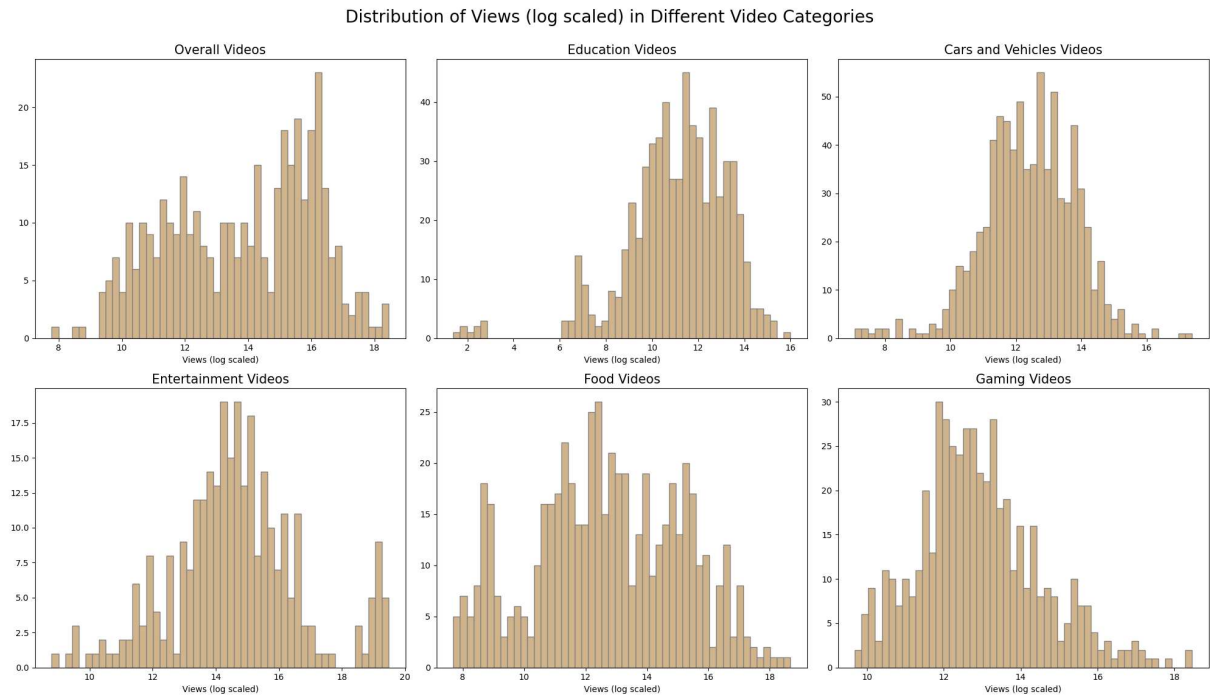
**Figure S1.** Distribution of number of views in the different video categories.



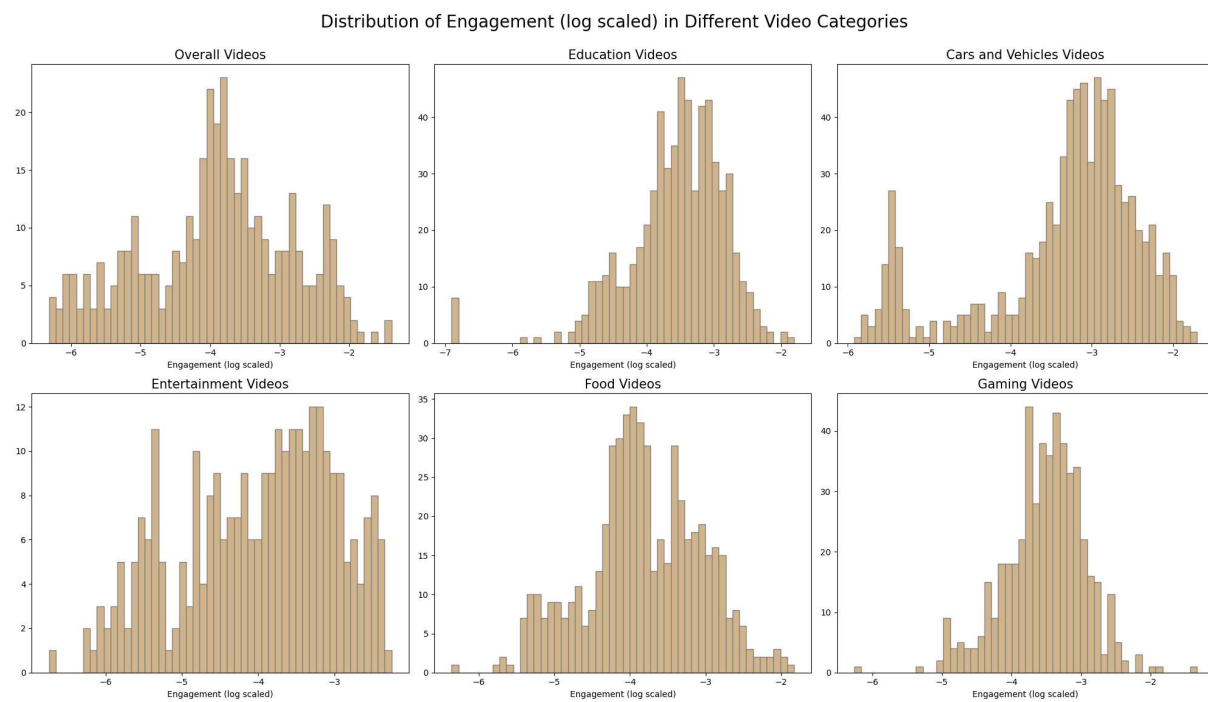
**Figure S2.** Distribution of engagement in the different video categories.



**Figure S3.** Distribution of video duration in the different video categories after log scaling.

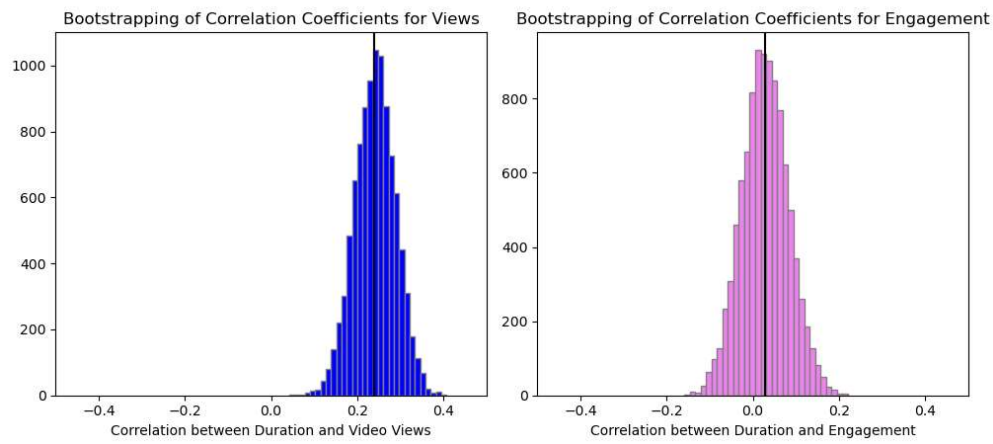


**Figure S4.** Distribution of number of views in the different video categories after log scaling.

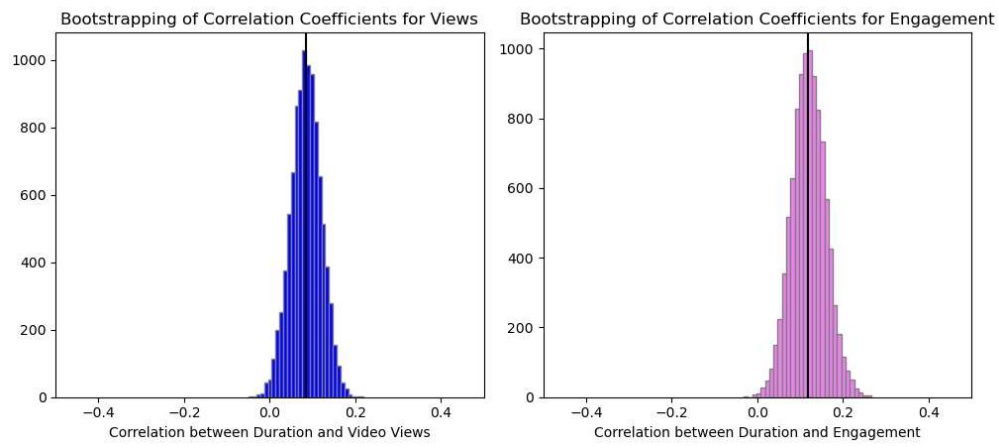


**Figure S5.** Distribution of engagement in the different video categories after log scaling.

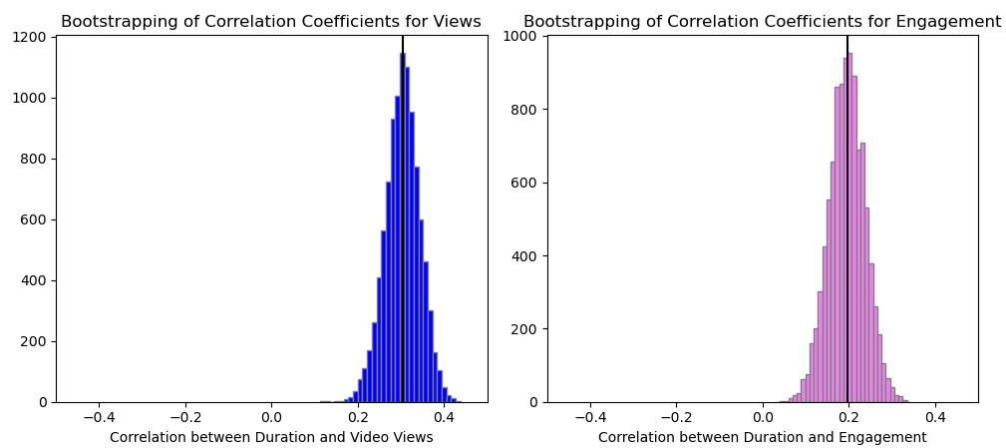
## Overall

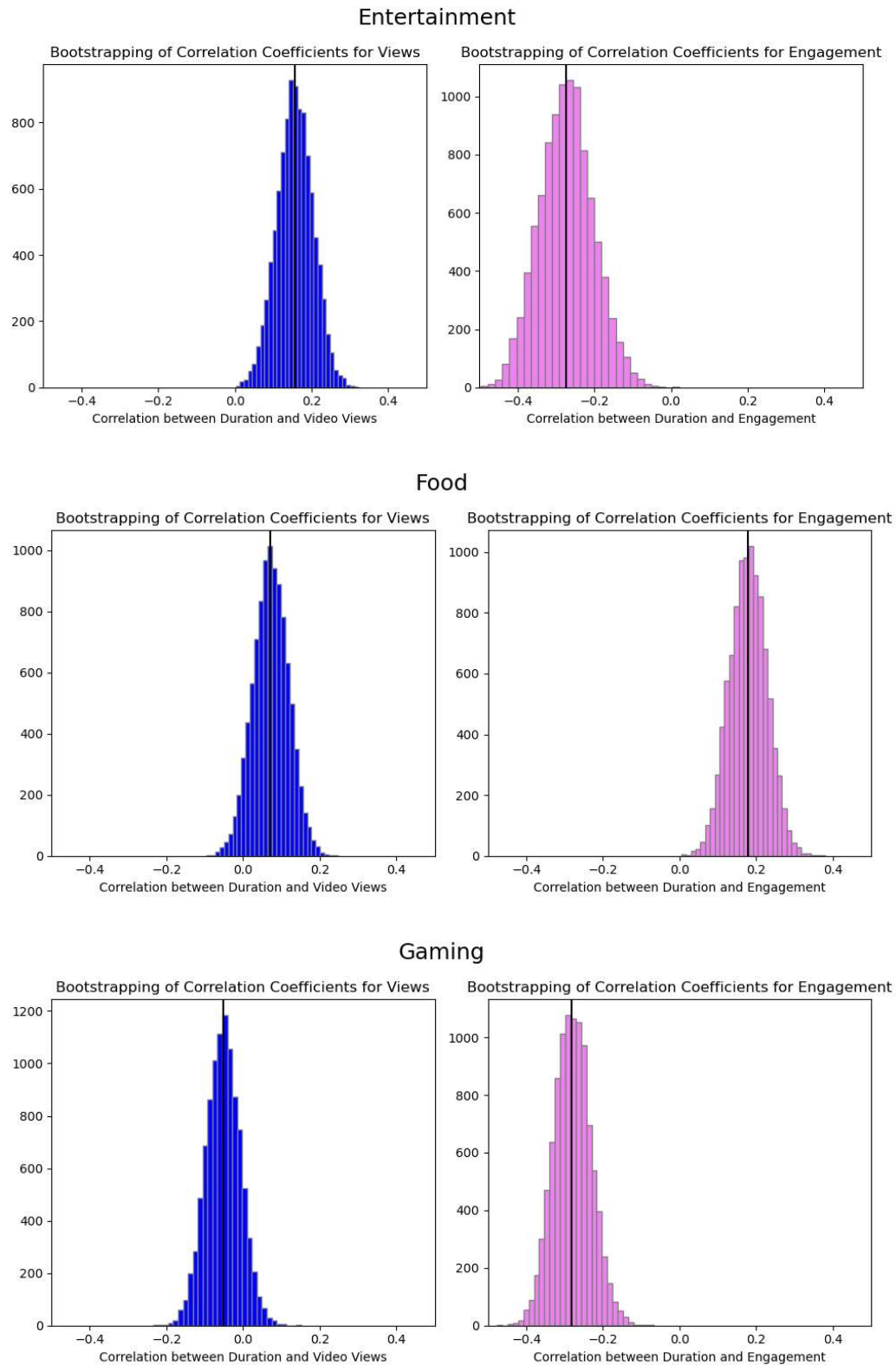


## Education



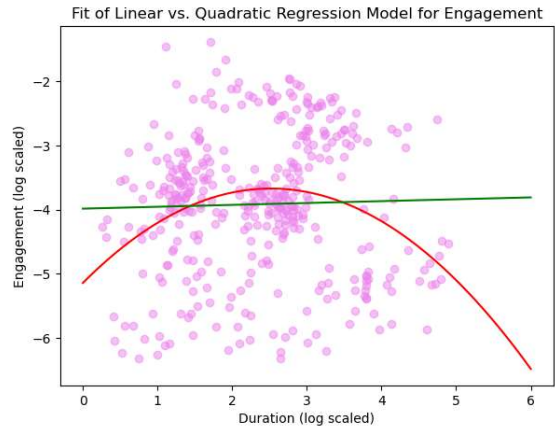
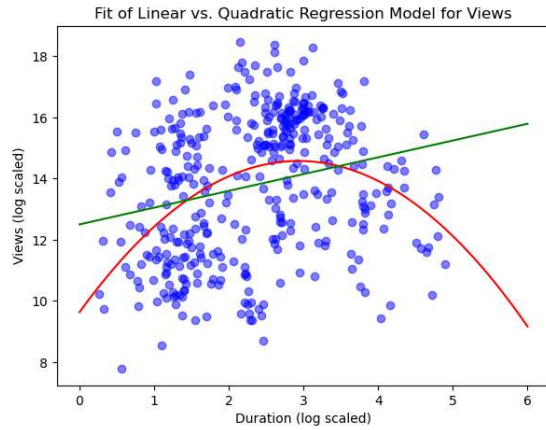
## Cars & Vehicles



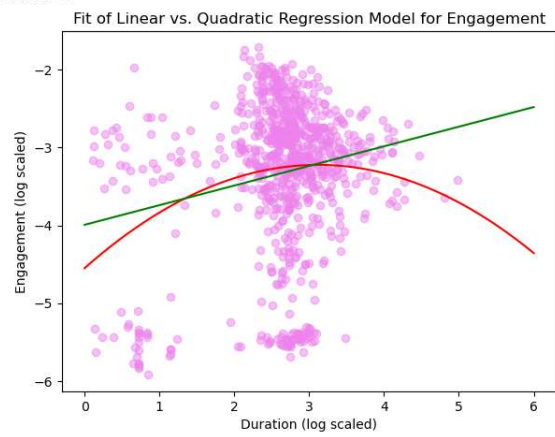
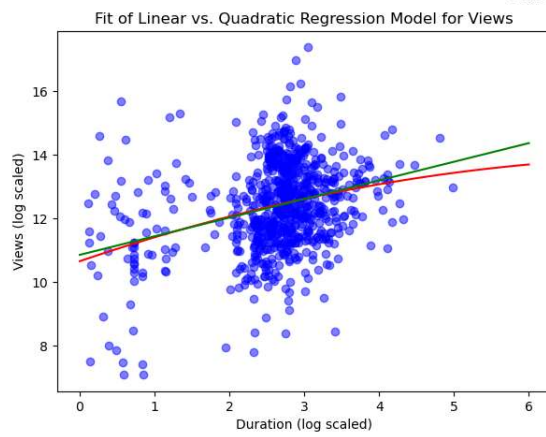


**Figure S6.** Distribution of the correlation coefficients from the 10,000 samples created by bootstrapping. One distribution for the correlation between duration and number of views (blue) and one for the correlation between duration and engagement (purple) for each of the six categories. The black line corresponds to the correlation in the original sample of videos.

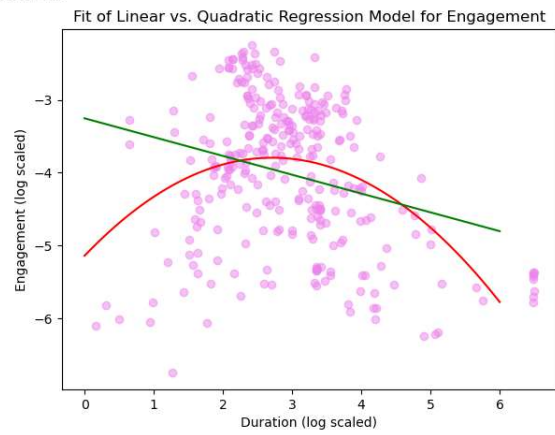
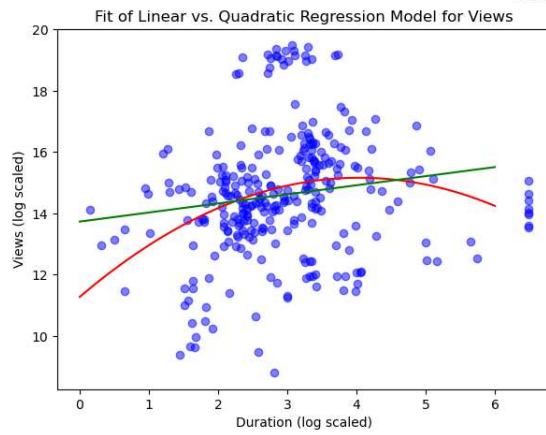
## Overall

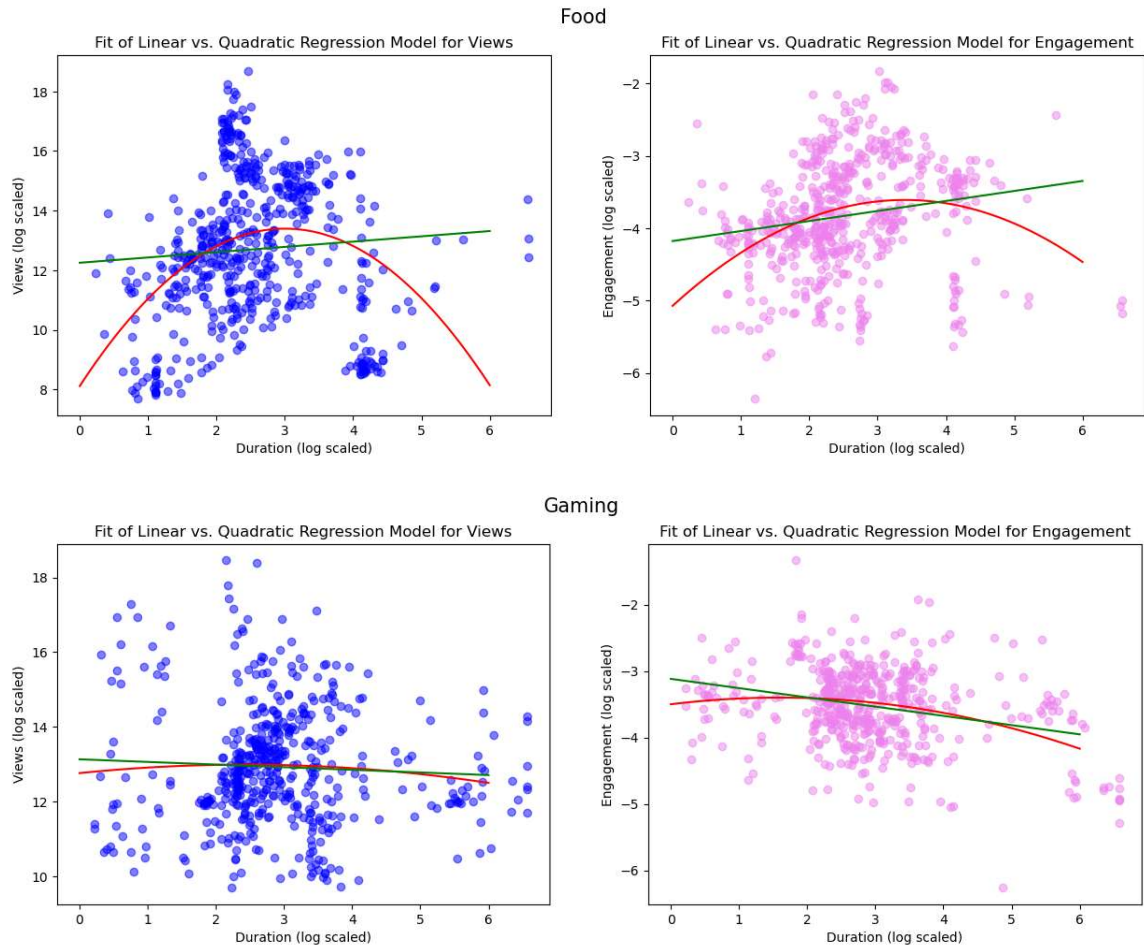


## Cars & Vehicles



## Entertainment





**Figure S7.** Scatterplots of the relationships between video duration and number of views (left) and video duration and engagement (right) for all categories. All variables are on a logarithmic scale. The green line corresponds to the fitted linear regression model, the red line to the quadratic regression model.

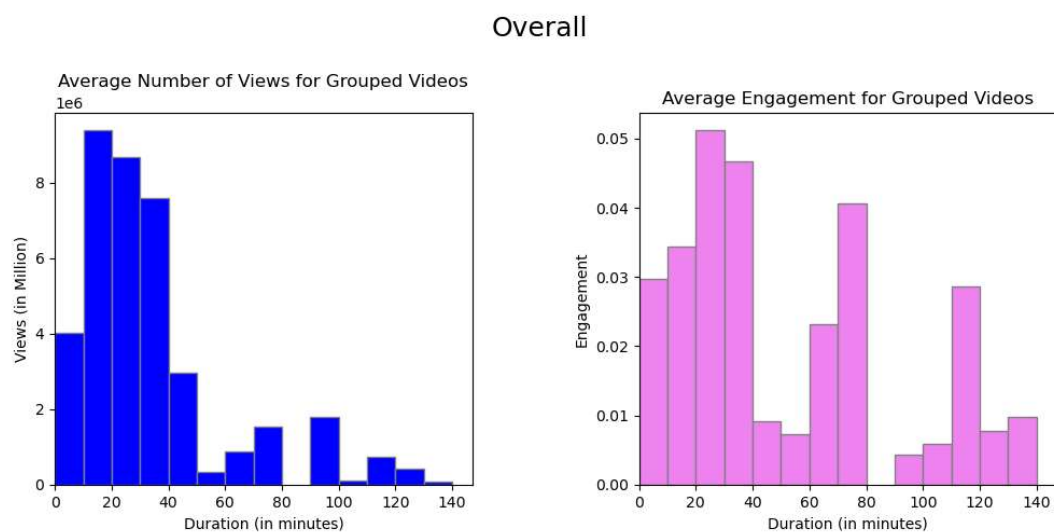
**Table S1.** Fit of linear vs. quadratic regression model for views as dependent variable

	$R^2$ (linear)	$R^2$ (quadratic)	F-statistic	p-value
Overall	0.058	0.147	39.757	< 0.001
Education	0.007	0.014	4.268	0.039
Cars & Vehicles	0.092	0.094	0.932	0.335
Entertainment	0.024	0.083	18.661	< 0.001
Food	0.005	0.144	88.519	< 0.001
Gaming	0.003	0.006	1.73	0.189

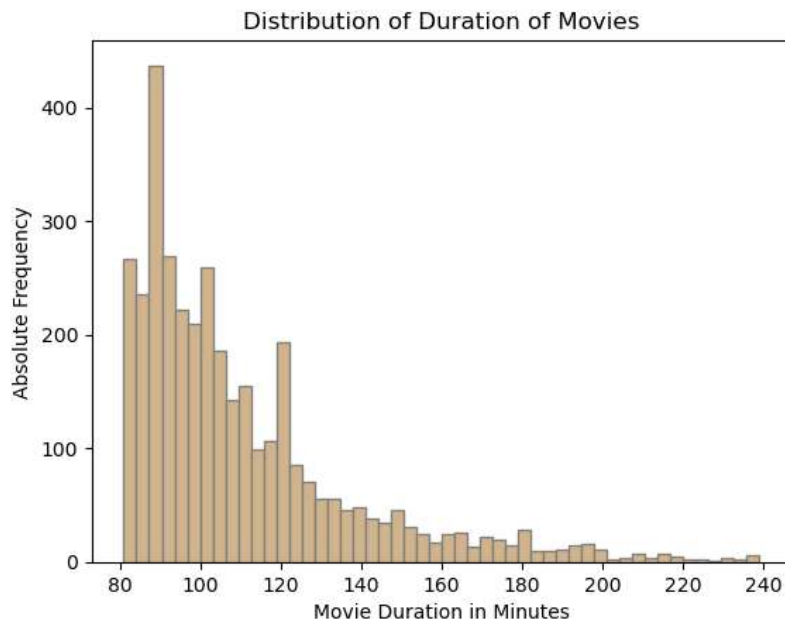


**Table S2.** Fit of linear vs. quadratic regression model for engagement as dependent variable

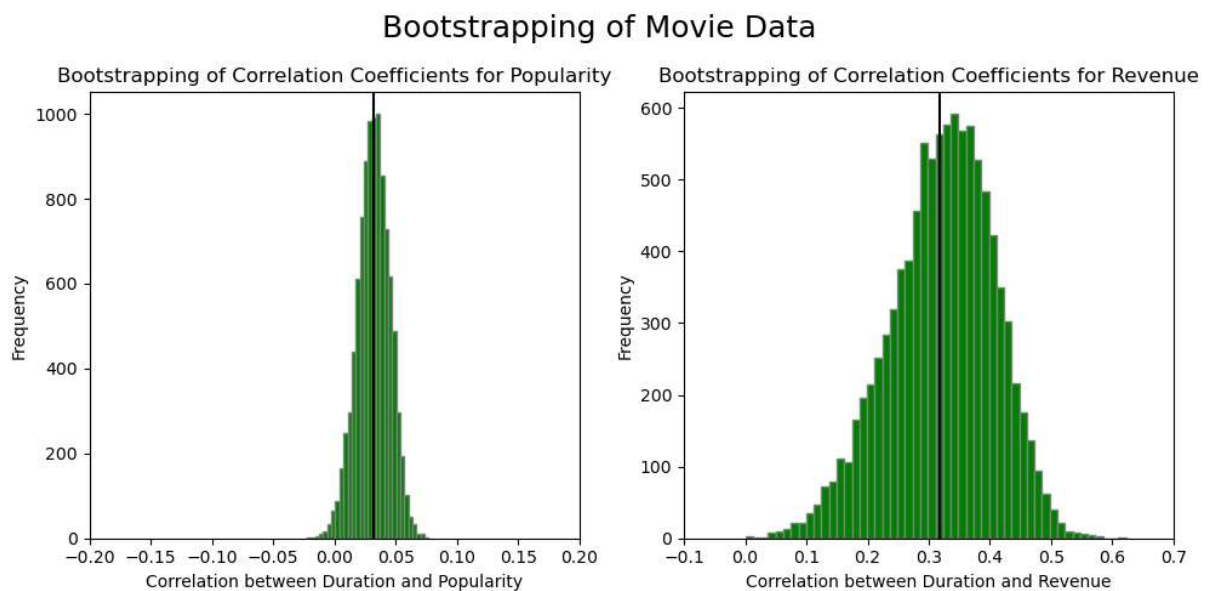
	$R^2$ (linear)	$R^2$ (quadratic)	F-statistic	p-value
Overall	0.001	0.067	26.965	< 0.001
Education	0.014	0.048	22.037	< 0.001
Cars & Vehicles	0.038	0.058	15.652	< 0.001
Entertainment	0.075	0.216	52.754	< 0.001
Food	0.032	0.097	39.613	< 0.001
Gaming	0.079	0.106	15.275	< 0.001



**Figure S8.** Videos of the overall category grouped by duration in steps of 10 minutes. For each group the mean number of views and the mean engagement is calculated and plotted. No bar means that no videos exist in this duration group.



**Figure S9.** Distribution of Movie Duration.



**Figure S10.** Distribution of the correlation coefficients from the 10,000 samples created by bootstrapping. One distribution for the correlation between duration and movie popularity (left) and one for the correlation between duration and revenue (right). The black line corresponds to the correlation in the original sample of movies.