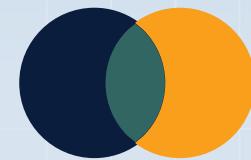


THE SKY'S THE LIMIT

ELEVATING LUFTHANSA'S INSTAGRAM MARKETING THROUGH DATA-DRIVEN INSIGHTS

Ballegeer Matteo	01900129
De Rocker Yaël	01904043
Janssens Wannes	01905583
Peire Julian	01900199
Willemen Tom	01900194





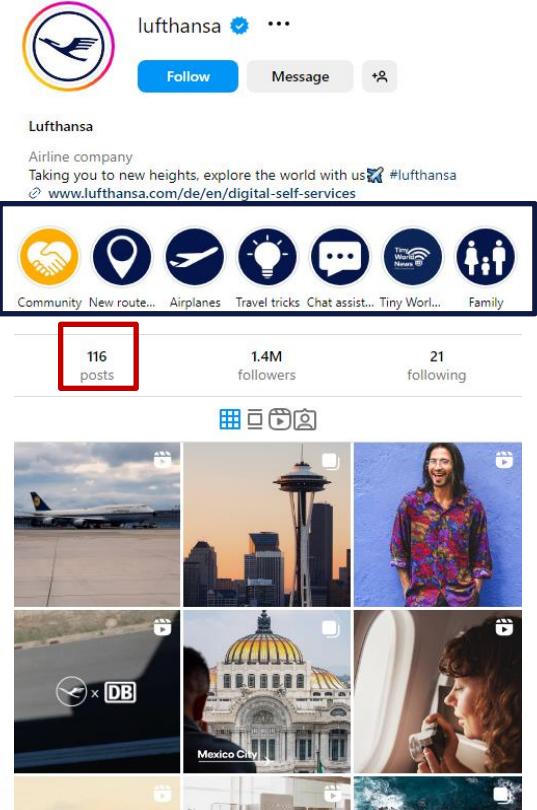
01

LUFTHANSA VS SINGAPOREAIR: A GENERAL COMPARISON

A QUALITATIVE COMPARISON

Story highlights

(Recently removed old posts)



Lufthansa

Airline company
Taking you to new heights, explore the world with us! #lufthansa
www.lufthansa.com/de/en/digital-self-services

Follow Message +2

Community New route... Airplanes Travel tricks Chat assist... Tiny Worl... Family

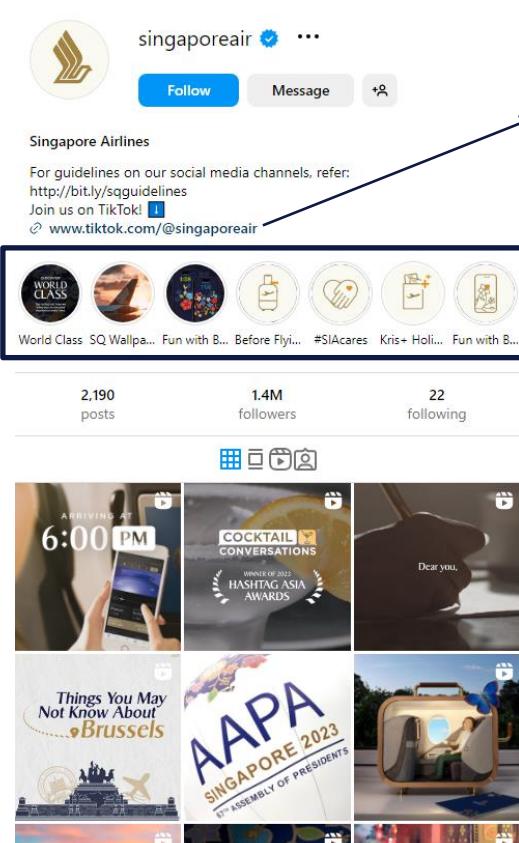
116 posts 1.4M followers 21 following

World Class SQ Wallpa... Fun with B... Before Flyi... #SIAcares Kris+ Hol... Fun with B...

Things You May Not Know About Brussels AAPA SINGAPORE 2023 6th ASSEMBLY OF PRESIDENTS

TikTok link

Story highlights



singaporeair

Singapore Airlines

For guidelines on our social media channels, refer:
<http://bit.ly/sqguidelines>
Join us on TikTok!
www.tiktok.com/@singaporeair

Follow Message +2

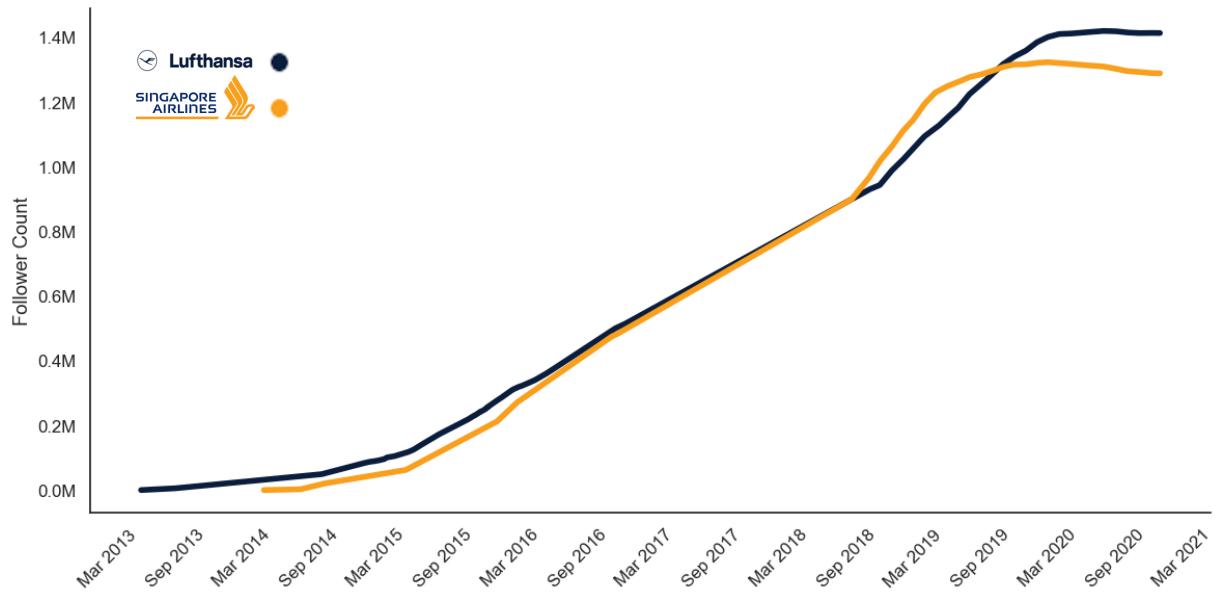
World Class SQ Wallpa... Fun with B... Before Flyi... #SIAcares Kris+ Hol... Fun with B...

2,190 posts 1.4M followers 22 following

AIRIVING AT 6:00 PM COCKTAIL CONVERSATIONS WINNER OF 2022 HASHTAG ASIA AWARDS Dear you,

BOTH COMPANIES SHARE A SIMILAR GROWTH STORY

Follower growth over time¹



2021 Industry comparison²

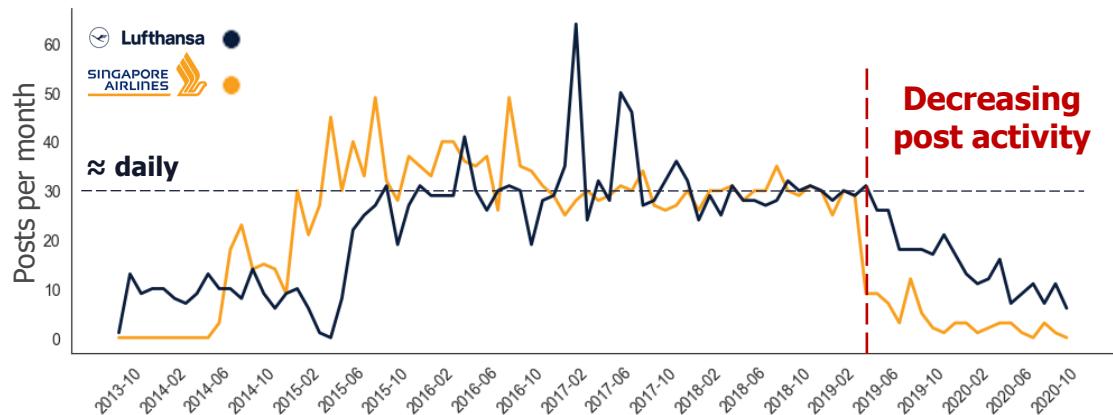
1. Emirates: 5.5M
2. Qatar Airways: 3.4M
3. LATAM Airlines: 1.8M
4. Turkish Airlines: 1.8M
5. Etihad: 1.7M
- 6. Lufthansa: 1.4M**
7. KLM: 1.3M
- 8. Singapore Air: 1.2M**
9. Garuda Indonesia: 1.1M
10. Air France: 1.0M

¹ Approximation using historical snapshots of follower counts: <https://web.archive.org/>

² <https://www.id1.de/2021/05/11/20-best-airline-brands-on-instagram-in-2021/>

INSTAGRAM POST ACTIVITY: A DECLINING TREND

Monthly post evolution



Macro factors

A shift in preference towards shorter, more personal video content

The Rapid Adoption of Instagram Stories



② The Oxford Blue

TikTok and the Death of the Attention Span

TikTok Was Installed More Than 738 Million Times in 2019, 44% of Its All-Time Downloads



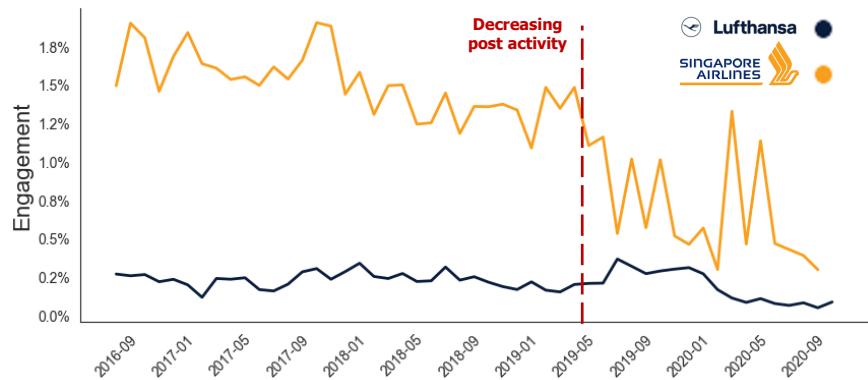
WHAT ABOUT ENGAGEMENT?

Definition¹

$$\frac{\text{likes} + \text{comments}}{\text{followers at time of posting}} * 100$$

= % of followers that liked or commented on a post

Monthly average



Average Engagement Benchmarks²

2019 (Overall average): 1.10%

2020 (Airline accounts): 2.27%

2021 (Travel accounts): 1.30%

2022 (Travel accounts): 0.91%

2023 (Aviation accounts): 1.07%

2020 Singapore Air: 0.60%

2020 Lufthansa: 0.13%

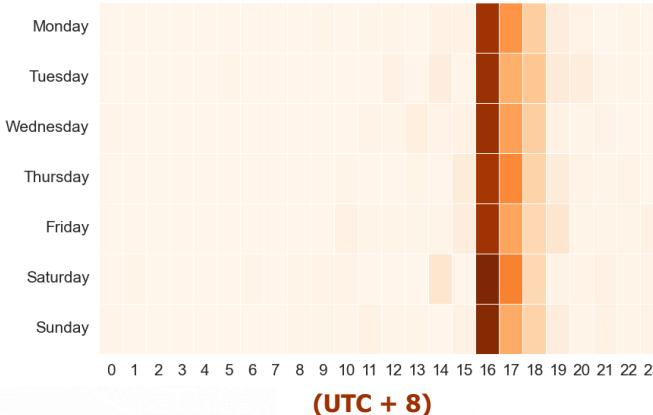
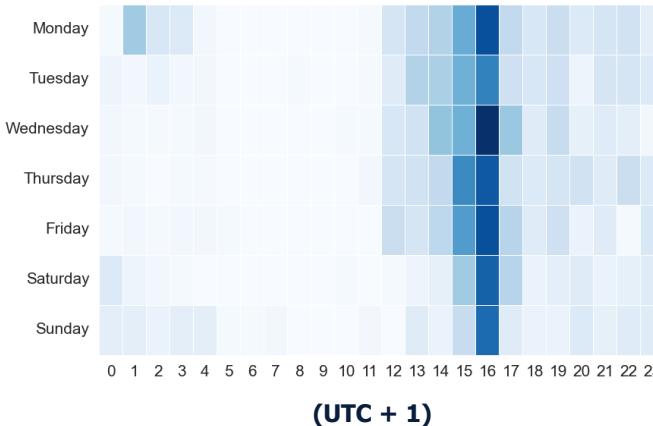
¹ Based on what is commonly found in literature

² Sources are appended in the *references* slide

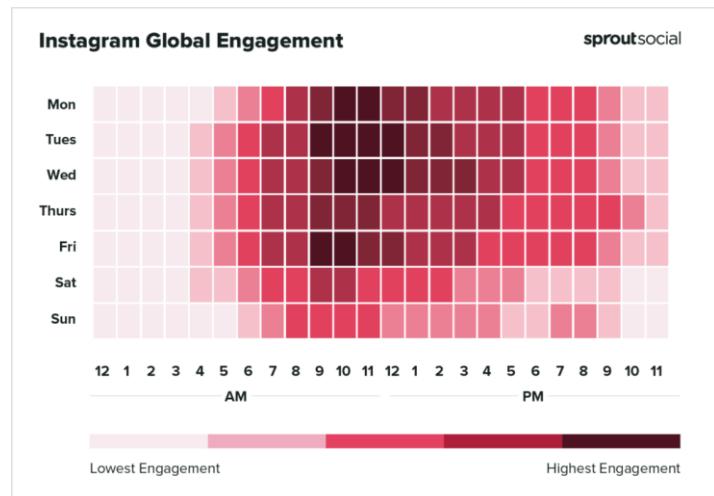


COMPARING BOTH AIRLINES' INSTAGRAM STRATEGY

DO WE OBSERVE DIFFERENCES IN POST TIMING?



**Observed engagement on over
34,000 Instagram accounts' posts¹**

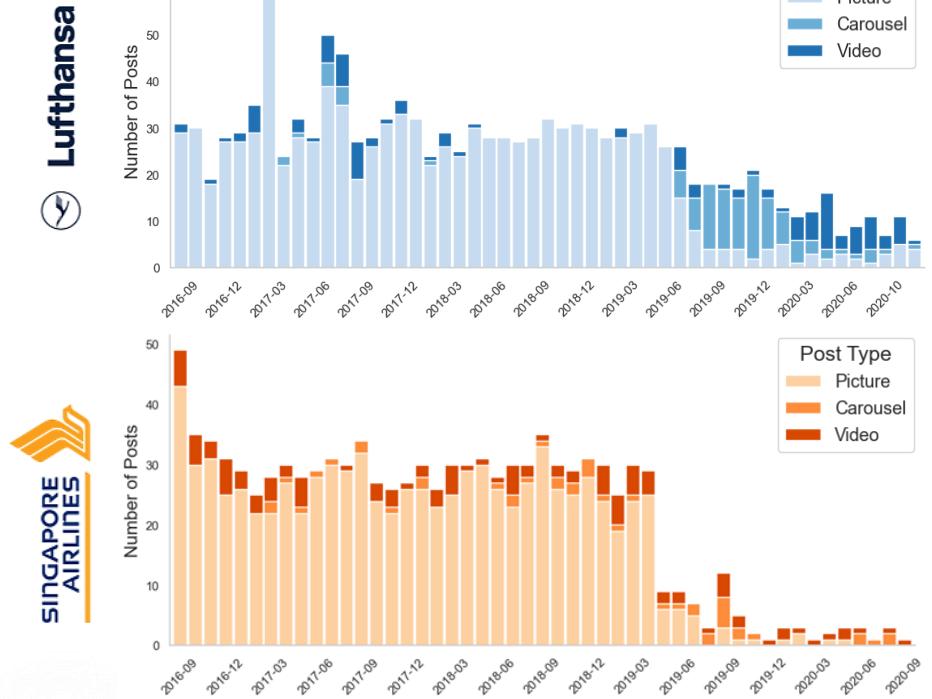


Focus on midweek mornings Avoid posting in weekends



¹ <https://sproutsocial.com/insights/best-times-to-post-on-social-media/>

LUFTHANSA IS MORE EXPERIMENTAL IN ITS POST TYPES



Other research on the topic¹

Table 4. Negative binomial regression results.

Parameter	Likes	
	B	Std. Error
(Intercept)	6.730**	.144
Photo (Base)	–	–
Short Video	-.012	.109
IGTV	-.139	.118
Carousel	.338**	.1

Average engagement rate per carousel, image, and video post²



Source: Socialinsider data
 Data range: January 2017 - July 2020

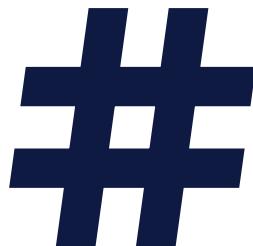
¹Risqo , W. M. (n.d.). Factors driving social media engagement on Instagram [Https://www.tandfonline.com/doi/full/10.1080/08911762.2021.1956665](https://www.tandfonline.com/doi/full/10.1080/08911762.2021.1956665)

²Socialinsider. (n.d.). Instagram Carousel Study - Socialinsider. [Https://www.socialinsider.io/data-geeks/Instagram-Carousel-Study.pdf](https://www.socialinsider.io/data-geeks/Instagram-Carousel-Study.pdf)

CAPTION STRUCTURE AND HASHTAGS



Lufthansa

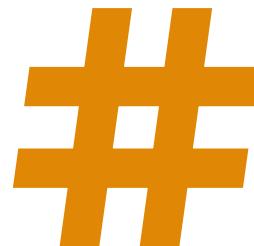


LhCityOfTheMonth
CityOfTheMonth
Heimweh
SayYesToTheWorld

58% Of posts contain hashtags



Travelling, exploring the world and its cultures



FlySQ
SQfoodie
SQavgeek
UltimateA350Experience

93% Of posts contain hashtags



Product quality, flight experience





WHAT ARE BOTH BRANDS TALKING ABOUT?

Lufthansa





SINGAPORE
AIRLINES



Adventure

Travel, place, view, world, city, see ...

Follower-taken photos

“Photo by ...”

Service & quality

**Book, board, experience,
meal, onboard, view, flight ...**

Prompts for comments

“Let us know ...”

“Share with us ...”



THIS DIFFERENCE IN FOCUS ALSO TRANSLATES INTO THE VISUALS



22%

Of posts contain
airplanes



Lufthansa

10%

Of posts show the
brand logo



33%

Of posts contain
airplanes



**SINGAPORE
AIRLINES**

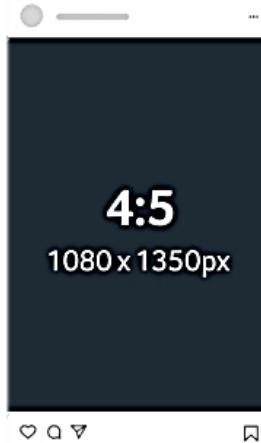


13%

Of posts show the
brand logo

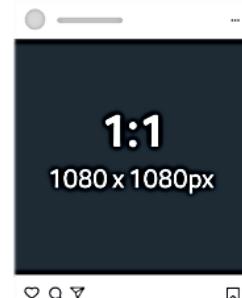


LUFTHANSA EXPERIMENTS MORE IN ITS ASPECT RATIO'S¹



36%

Portrait



40%

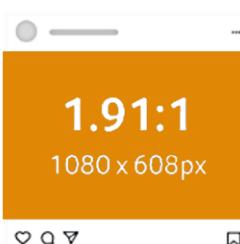
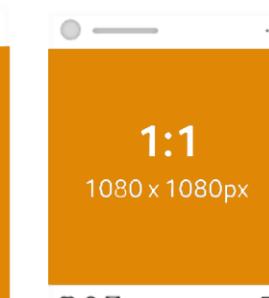


24%
Landscape



18%

Portrait



29%
Landscape



¹ Kastrenakes, J. (2015, September 27). *Portrait is the new square on Instagram*. The Verge. <https://www.theverge.com/2015/9/27/9398551/instagram-portrait-photos-rule>

LUFTHANSA LEVERAGES TAGGING MORE OFTEN



Lufthansa



6.44%

of posts **explicitly**
tag other accounts

Of which

18%

are **verified**



63%

of posts contain
tags through an
@ in the caption



56%

of posts
tag a city

Of which only

22%

in Germany



0.62%

of posts **explicitly**
tag other accounts

Of which

25%

are **verified**



68%

of posts contain
tags through an
@ in the caption



2%

of posts
tag a city

Of which

93%

in Singapore



03

IDENTIFYING THE DRIVERS OF INSTAGRAM ENGAGEMENT

POTENTIAL DRIVERS

Insight-Derived

Literature-Based



Caption



Contains City



Nr. @



Nr. #'¹



Nr. '?'¹



Nr. '!'¹



Contains Link¹



Sentiment²



Nr.Words¹



Nr.Tags



Subjectivity²



Nr. Emojis



Post



Part of Day



Season



Time Top Post¹



Is Weekday¹



Location



Aspect Ratio³



Carousel



Vividness¹



Image



Contains Airplane



Lufthansa



Contains Logo



Nr. Faces⁴



Colourfulness¹



Contains Text

¹ De Vries, L., Gensler, S., & Leeflang, P. S. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of interactive marketing*, 26(2), 83-91.

² Meire, M., Hewett, K., Ballings, M., Kumar, V., & Van den Poel, D. (2019). The role of marketer-generated content in customer engagement marketing. *Journal of Marketing*, 83(6), 21-42.

³ Kastrenakes, J. (2015, September 27). *Portrait is the new square on Instagram*. The Verge.
<https://www.theverge.com/2015/9/27/9398551/instagram-portrait-photos-rule>

⁴ Öhman, W. (2017). *Instagram marketing: a study about the effect of visual content on customer engagement in the airline industry* (Doctoral dissertation).



POTENTIAL DRIVERS

Insight-Derived

Literature-Based



Caption



Contains City



Nr. @



Nr. '#'



Nr. '?'



Nr. 'I'


 Contains Link¹

 Sentiment²

 Nr.Words¹


Nr.Tags


 Subjectivity²


Nr.Emojis



Post



Part of Day



Season


 Time Top Post¹

 Is Weekday¹


Location


 Aspect Ratio³


Carousel


 Vividness¹


Image



Contains Airplane



Contains Logo


 Nr. Faces⁴

 Colourfulness¹


Contains Text

¹De Vries, L., Gensler, S., & Leeftlang, P. S. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of interactive marketing*, 26(2), 83-91.

²Meire, M., Hewett, K., Ballings, M., Kumar, V., & Van den Poel, D. (2019). The role of marketer-generated content in customer engagement marketing. *Journal of Marketing*, 83(6), 21-42.

³Kastrenakes, J. (2015, September 27). *Portrait is the new square on Instagram*. The Verge.
<https://www.theverge.com/2015/9/27/9398551/instagram-portrait-photos-rule>

⁴Öhman, W. (2017). *Instagram marketing: a study about the effect of visual content on customer engagement in the airline industry* (Doctoral dissertation).



POTENTIAL DRIVERS

Insight-Derived

Literature-Based



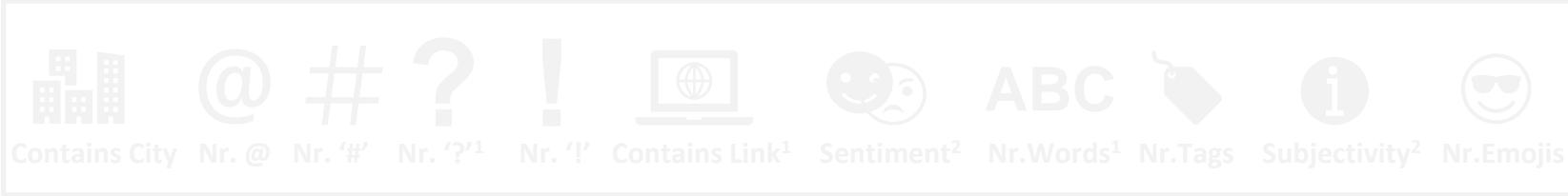
Caption



Post



Image



¹De Vries, L., Gensler, S., & Leeflang, P. S. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of interactive marketing*, 26(2), 83-91.

²Meire, M., Hewett, K., Ballings, M., Kumar, V., & Van den Poel, D. (2019). The role of marketer-generated content in customer engagement marketing. *Journal of Marketing*, 83(6), 21-42.

³Kastrenakes, J. (2015, September 27). *Portrait is the new square on Instagram*. The Verge.
<https://www.theverge.com/2015/9/27/9398551/instagram-portrait-photos-rule>

⁴Öhman, W. (2017). *Instagram marketing: a study about the effect of visual content on customer engagement in the airline industry* (Doctoral dissertation).



POTENTIAL DRIVERS

Insight-Derived

Literature-Based



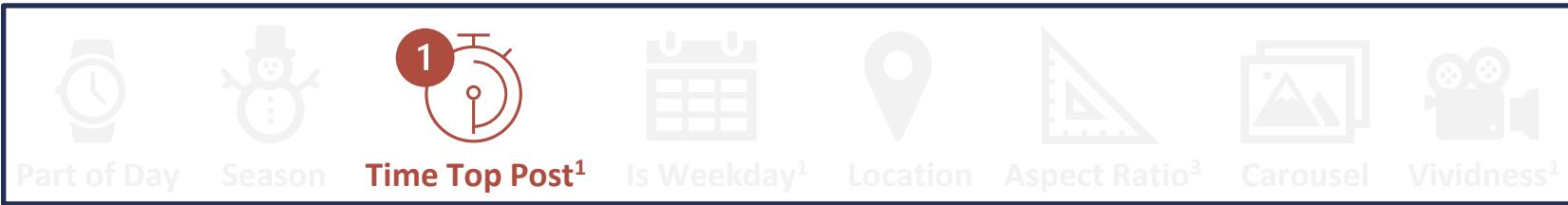
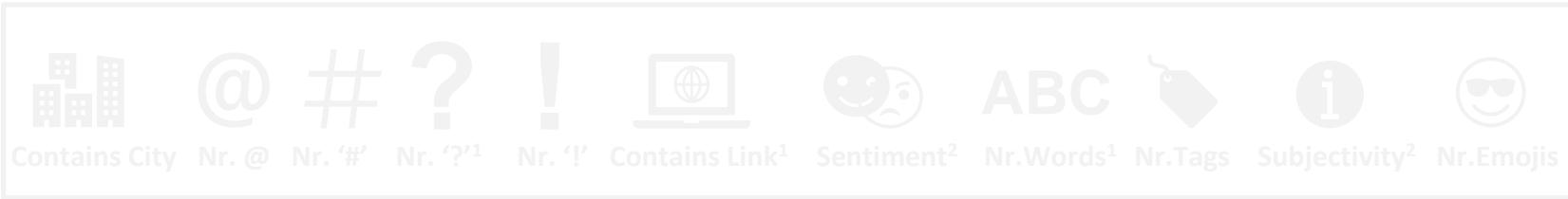
Caption



Post



Image



¹De Vries, L., Gensler, S., & Leeflang, P. S. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of interactive marketing*, 26(2), 83-91.

²Meire, M., Hewett, K., Ballings, M., Kumar, V., & Van den Poel, D. (2019). The role of marketer-generated content in customer engagement marketing. *Journal of Marketing*, 83(6), 21-42.

³Kastrenakes, J. (2015, September 27). *Portrait is the new square on Instagram*. The Verge.
<https://www.theverge.com/2015/9/27/9398551/instagram-portrait-photos-rule>

⁴Öhman, W. (2017). *Instagram marketing: a study about the effect of visual content on customer engagement in the airline industry* (Doctoral dissertation).



POTENTIAL DRIVERS

Insight-Derived

Literature-Based



Caption



Post



Image



¹ De Vries, L., Gensler, S., & Leeflang, P. S. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of interactive marketing*, 26(2), 83-91.

² Meire, M., Hewett, K., Ballings, M., Kumar, V., & Van den Poel, D. (2019). The role of marketer-generated content in customer engagement marketing. *Journal of Marketing*, 83(6), 21-42.

³ Kastrenakes, J. (2015, September 27). *Portrait is the new square on Instagram*. The Verge.
<https://www.theverge.com/2015/9/27/9398551/instagram-portrait-photos-rule>

⁴ Öhman, W. (2017). *Instagram marketing: a study about the effect of visual content on customer engagement in the airline industry* (Doctoral dissertation).



POTENTIAL DRIVERS

Insight-Derived

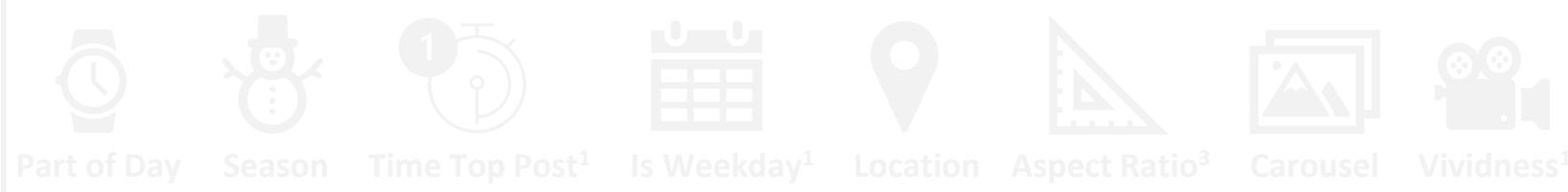
Literature-Based



Caption



Post



Image



¹ De Vries, L., Gensler, S., & Leeflang, P. S. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of interactive marketing*, 26(2), 83-91.

² Meire, M., Hewett, K., Ballings, M., Kumar, V., & Van den Poel, D. (2019). The role of marketer-generated content in customer engagement marketing. *Journal of Marketing*, 83(6), 21-42.

³ Kastrenakes, J. (2015, September 27). *Portrait is the new square on Instagram*. The Verge.
<https://www.theverge.com/2015/9/27/9398551/instagram-portrait-photos-rule>

⁴ Öhman, W. (2017). *Instagram marketing: a study about the effect of visual content on customer engagement in the airline industry* (Doctoral dissertation).



POTENTIAL DRIVERS

Insight-Derived

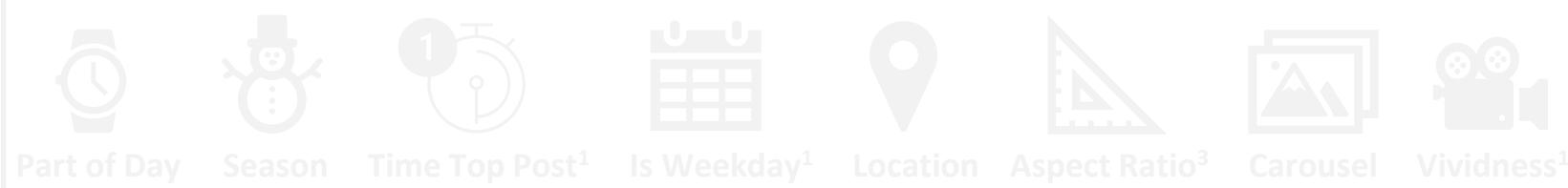
Literature-Based



Caption



Post



Image



¹ De Vries, L., Gensler, S., & Leeflang, P. S. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of interactive marketing*, 26(2), 83-91.

² Meire, M., Hewett, K., Ballings, M., Kumar, V., & Van den Poel, D. (2019). The role of marketer-generated content in customer engagement marketing. *Journal of Marketing*, 83(6), 21-42.

³ Kastrenakes, J. (2015, September 27). *Portrait is the new square on Instagram*. The Verge.
<https://www.theverge.com/2015/9/27/9398551/instagram-portrait-photos-rule>

⁴ Öhman, W. (2017). *Instagram marketing: a study about the effect of visual content on customer engagement in the airline industry* (Doctoral dissertation).



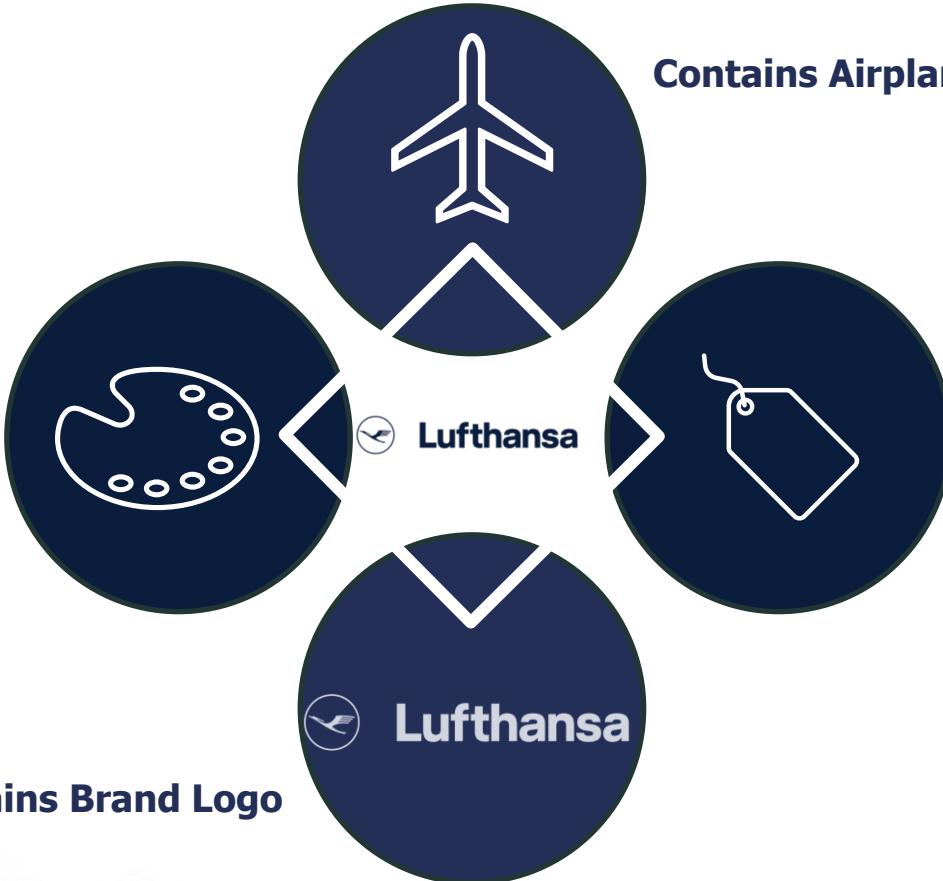
DRIVERS

Colourfulness

Contains Airplane

Number of Tags

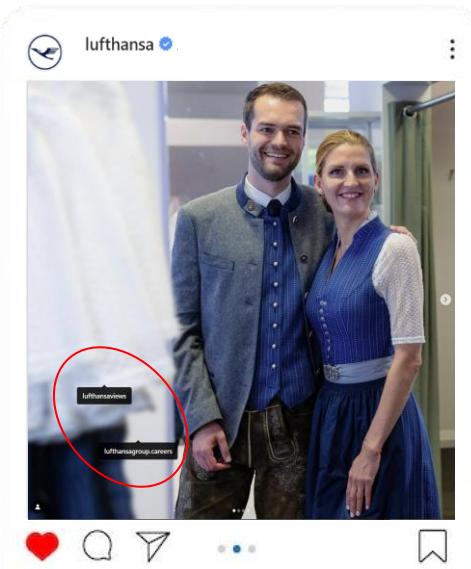
Contains Brand Logo



OUR RECOMMENDATIONS



POST MORE PICTURES
WITH A VISIBLE BRAND
LOGO AND PLANES



REMEMBER TO TAG USERS TO
HIGHLIGHT THEIR WORK



INCORPORATE LESS COLOUR
INTO YOUR PHOTOS NOT TO
OVERWHELM YOUR AUDIENCE



FURTHER STEPS

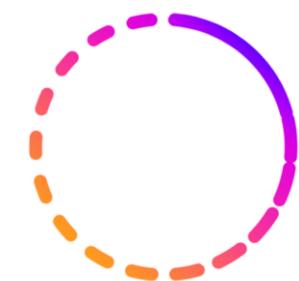
New forms of social media

Provide us with data regarding

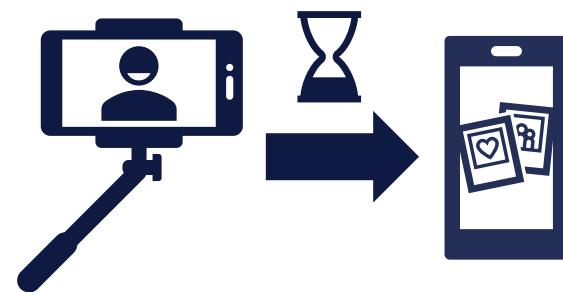
TikTok

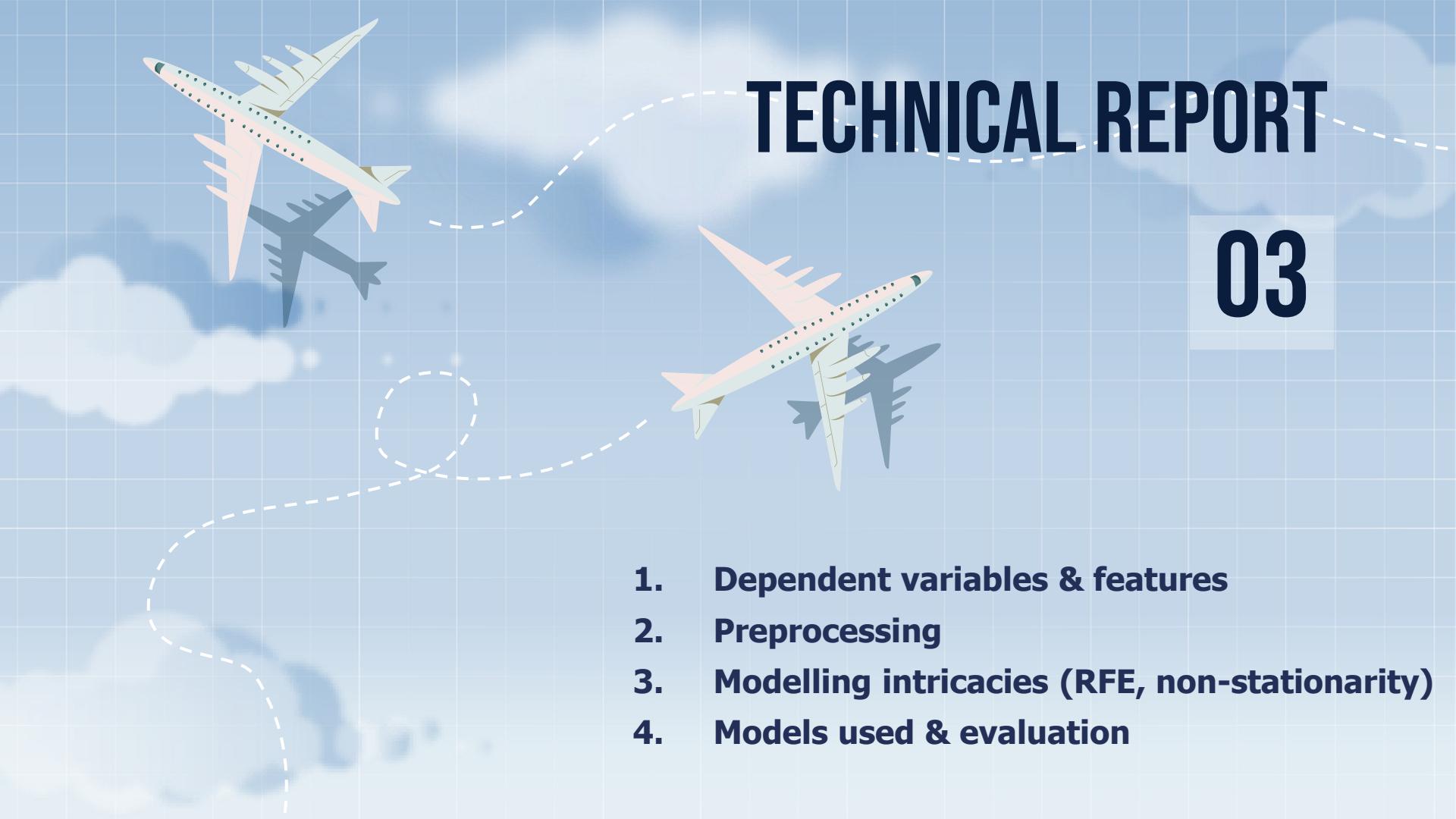
Instagram Stories

Instagram Reels



Analyse impact influencer posts





TECHNICAL REPORT

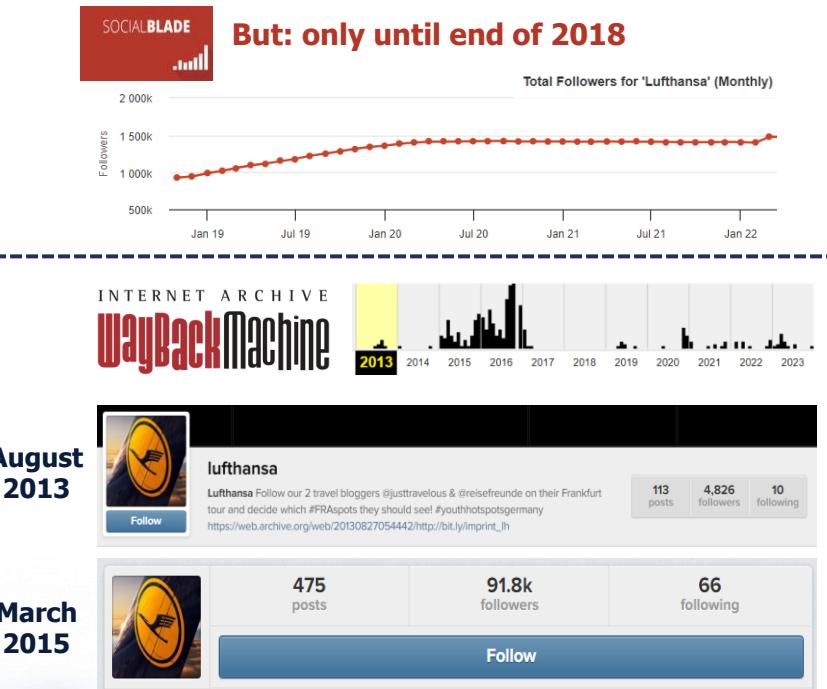
03

1. **Dependent variables & features**
2. **Preprocessing**
3. **Modelling intricacies (RFE, non-stationarity)**
4. **Models used & evaluation**

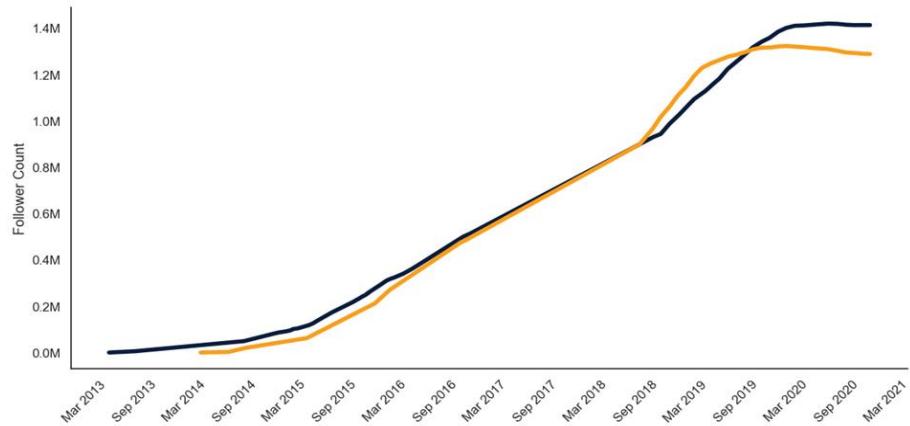
NUMBER OF FOLLOWERS

$$\text{Engagement} = \frac{\text{likes} + \text{comments}}{\text{followers at time of posting}} * 100$$

Problem: followers at time of posting = followers at time of scraping = constant
→ We need follower data from March 2013 – December 2020



Followers over time



FEATURE: HAS_BRAND_LOGO

Reasoning: Including a brand logo can signify the main subject of the content, potentially suggesting a commercial intent behind the post.

Step 1: Google Vision API — Logo Detection

Step 2: Filter

```
+-----+-----+
|Description |count|
+-----+-----+
|Lufthansa |9716 |
|Star Alliance |433 |
|Singapore Airlines |377 |
|Lufthansa Cargo |240 |
|British Airways |96 |
|LSG Sky Chefs |86 |
|United Airlines |81 |
|GeminiJets |79 |
|Emirates |76 |
|Lufthansa Systems |54 |
|Korean Air |44 |
|Air France |43 |
|Alaska Airlines |35 |
|Boeing |33 |
|Qatar Airways |32 |
|American Airlines |32 |
|LIU Post |31 |
|Swiss International Air Lines |27 |
|Monetary Authority of Singapore|25 |
|Qantas |25 |
+-----+
only showing top 20 rows
```

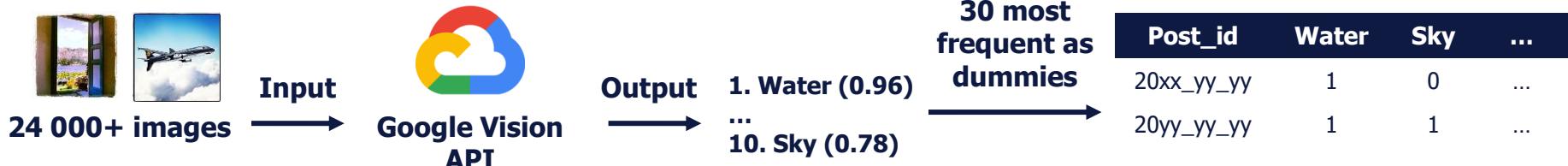
```
# Filter the DataFrame to only keep relevant logos
filtered_df = logo_df.filter(
    (col("Description").like("%Lufthansa%")) |
    (col("Description").like("%Singapore Airlines%")) |
    (col("Description").like("%Star Alliance%")))\n    .dropDuplicates(["name"])
```

Step 3: has_brand_logo feature
if picture has at least one of the
specified Logo's

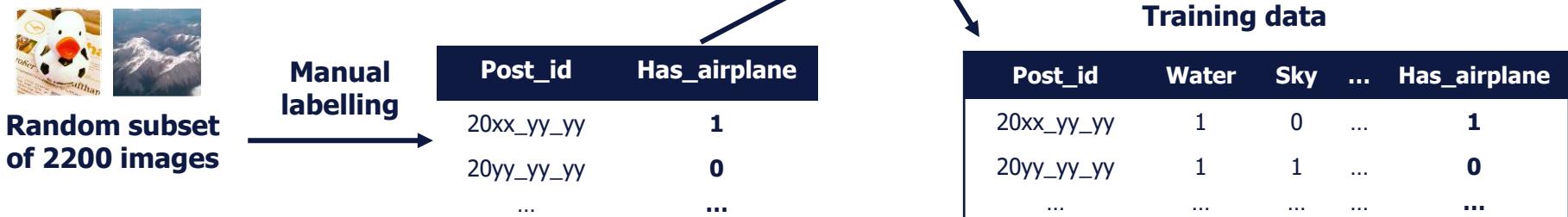


FEATURES: HAS_AIRPLANE (PART1)

Step 1: Use Google Vision API — Generate Descriptions



Step 2: Construct training data



Step 3: Create classification model that predicts has_airplane based on descriptions

- Train-validation-test split 0.56-0.24-0.20
- Random Forest — Decision Tree — Logistic Regression

FEATURES: HAS_AIRPLANE (PART 2)

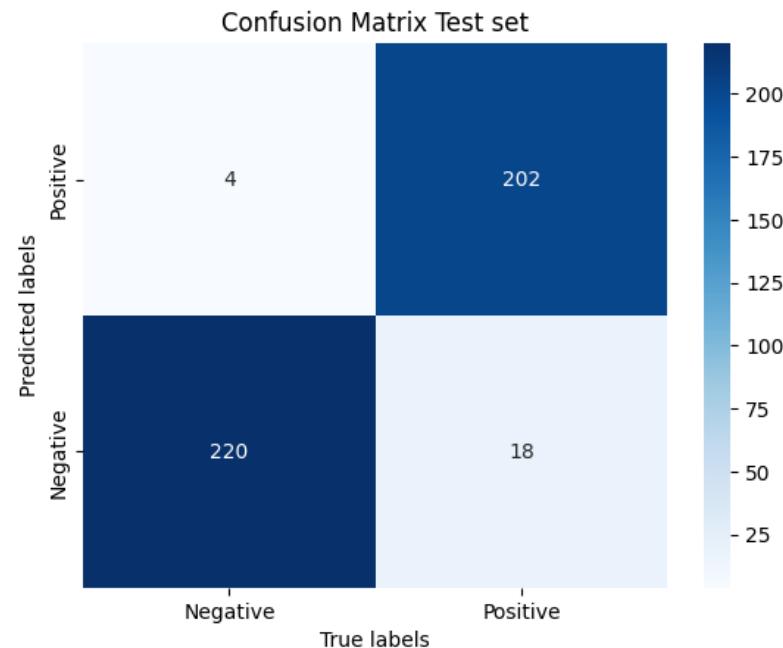
Step 4: Deduct best alternative on validation set — Logistic Regression with 25 most frequent descriptions

Step 5: Evaluation of performance

	Validation set	Test set
Accuracy	0.9728	0.9807
Precision	0.9569	0.9806
Recall	0.9136	0.9181

Step 6: Apply on remaining data

Step 7: Use as feature in engagement model



FEATURES: NR_FACES¹

Reasoning: Including people in pictures fosters human connection and relatability, utilizing the emotional impact of facial expressions to boost engagement by resonating with viewers' emotions.

Not perfect!

Face Count: 1



Face Count: 39



Face Count: 1



¹Using MTCNN package in Python: <https://pypi.org/project/mtcnn/>

Bakhshi, S., Shamma, D. A., & Gilbert, E. (2014). Faces engage us: photos with faces attract more likes and comments on Instagram. CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 965–974. <https://doi.org/10.1145/2556288.2557403>

FEATURES: CONTAINS_CITY

Reasoning: Lufthansa often mentions cities and uses tags such as #lhcityofthemonth. Chauvinism of followers of the mentioned region might generate a higher engagement

Method 1: Matching with external database



simplemaps 44k+
cities
Interactive Maps & Data

```
# Join DataFrame with cities DataFrame
contains_city = basetable_tokenized.join(
    cities_df,
    expr("lower(token) = lower(city)"),
    "left_outer")\
    .withColumn("contains_city_joins", col("city").isNotNull().cast("boolean"))\
    .groupBy("post_id", "cleaned_text")\
    .agg(expr("max(contains_city_joins) as contains_city_joins"))\
    .select("post_id", "cleaned_text", "contains_city_joins")

|contains_city_joins|count|
+-----+-----+
|      true| 1013|
|     false|  748|
+-----+-----+
```

Method 2: Named Entity Recognition (NER)

```
: ## An example of the NER at work
displacy.render(NER(contains_city.select("cleaned_text").collect()[3]["cleaned_text"]), style="ent")
displacy.render(NER(contains_city.select("cleaned_text").collect()[19]["cleaned_text"]), style="ent")
displacy.render(NER(contains_city.select("cleaned_text").collect()[17]["cleaned_text"]), style="ent")
```

Flying over Miami GPE lufthansa Thanks to tnrd for sharing

Some traveltuesday inspiration from beautiful Marseille France ORG How many of our followers have been there

traveltuesday DATE Summer DATE in Greece GPE Who wants to go

```
+-----+-----+
|contains_city_NER|count|
+-----+-----+
|      true| 348|
|     false| 1413|
+-----+-----+
```

```
+-----+-----+
|contains_city|count|
+-----+-----+
|      true| 331|
|     false| 1430|
+-----+-----+
```



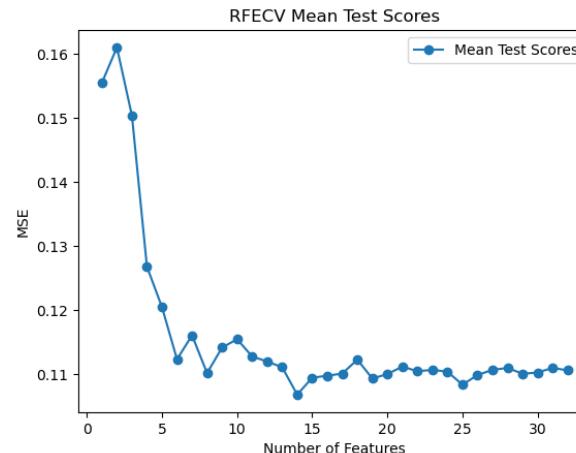
FEATURE SELECTION

Recursive Feature Elimination (RFE)

- No support for **PySpark** → Implemented through **scikit** (Python)

```
# Create the model for RFE
model = RandomForestRegressor()

min_features_to_select = 1
# Create the RFECV object
rfecv = RFECV(
    estimator=model,
    step=1,
    min_features_to_select=min_features_to_select,
    cv=15,
    scoring='neg_mean_squared_error',
    verbose=1,
    importance_getter="feature_importances_"
)
# Fit the RFECV object to the data
rfecv.fit(X_train, y_train)
```



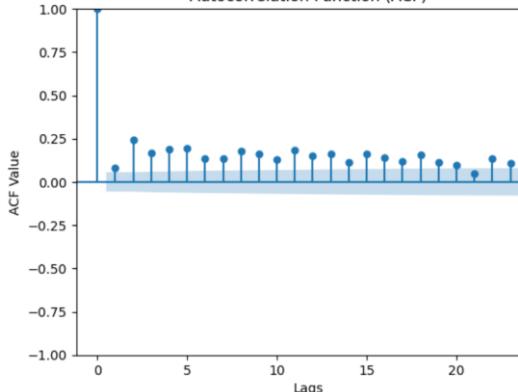
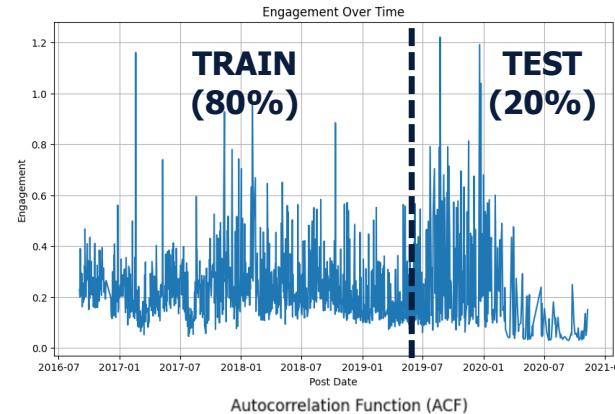
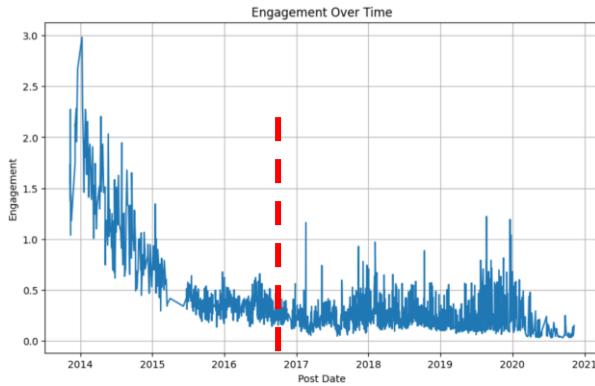
14 variables

- BUT:** Selection did not seem robust (different variables each time)
- **Discontinued feature selection**



DATA CUT-OFF

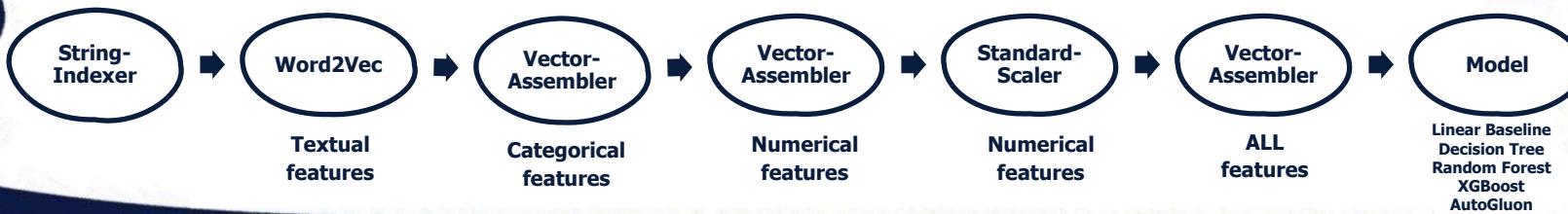
Cut-off in September 2016: data before it clearly shows trend and makes data unstationary



Autocorrelation better, but let's still perform a time-based split



PIPELINE SETUP



HYPERPARAMETER TUNING

Linear Regression

Regparam
[0,0.1,0.05,**0.01**,0.005,0.001]

ElasticNet
[0, 0.5, 1]

Decision Tree

MaxDepth
[**3**,5,8,15,20,30]

MinInstancesPerNode
[1, 5, 10, 25,**50**, 100]

MinInfoGain
[0, 0.1]
MaxBins
[**16**,32]

Random Forest

MaxDepth
[2,3,5,**7**]
MinInstancesPerNode

[1,5,**10**]
featureSubsetStrategy

[**p/3**]
subsamplingRate
[0.5,0.75,**1**]
numTrees
[**150**]

Gradient Boosting

MaxDepth
[**2**,4,6]
MaxIter
[10,20,30,**50**,100]

stepSize
[0.1]
MaxBins
[**32**]



MODEL COMPARISON

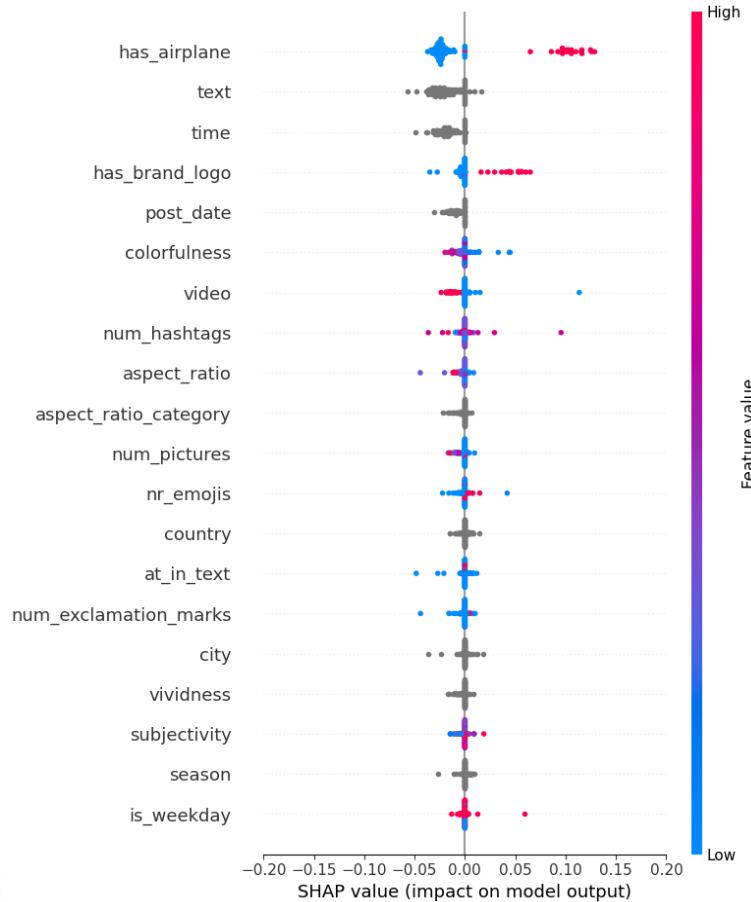
Based on cross-validation score

Model	RMSE
Linear Regression	0.1112
Decision Tree	0.0953
Gradient Boosting	0.0933
Random Forest	0.0911
AutoGluon	0.0817

Best model score on test set

Model	RMSE	MSE	MAE	R ²
AutoGluon	0.1600	0.0256	0.0978	0.4405

SHAP – AUTOGLUON



FURTHER CONSIDERATIONS / IMPROVEMENTS

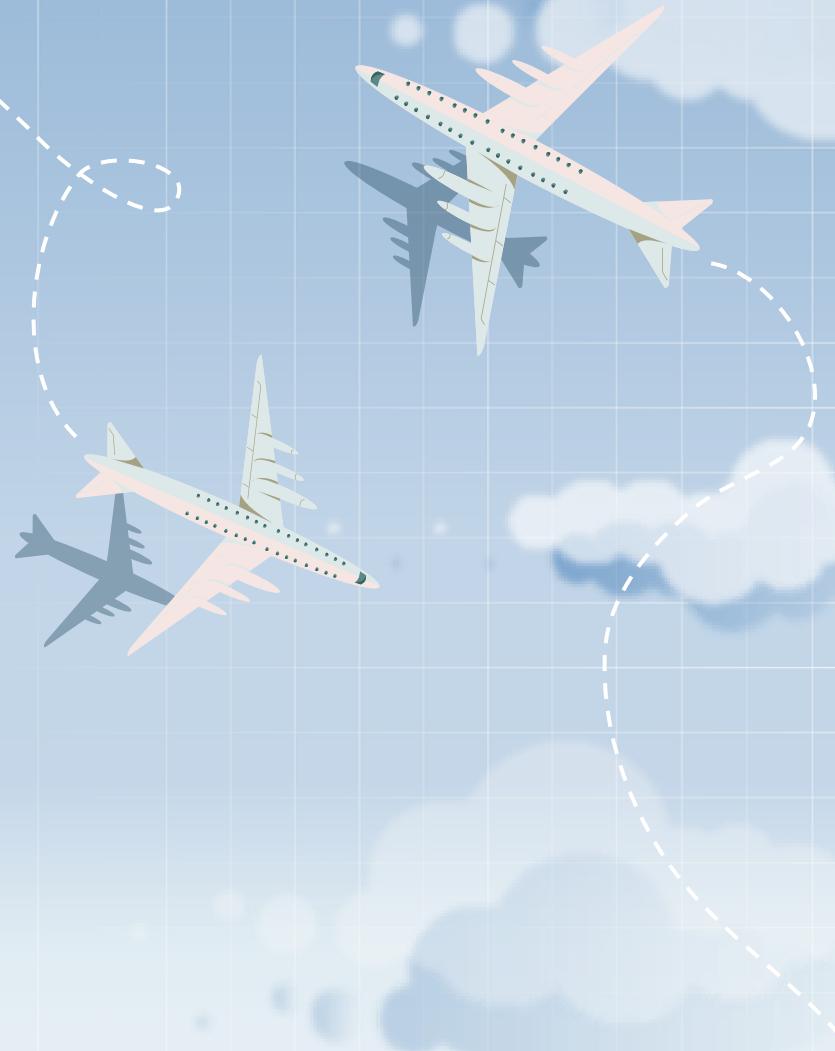
- 1. Engagement rate calculations currently include only initial comments**
 - Replies were not included, this should be done to be fully correct
 - On average this raises the number of comments by 8.3%
 - **BUT:** after deducting company responses (not considered follower engagement), only 6.4%
- 2. Still some autocorrelation / non-stationarity in the data when modelling**
 - Make scores on the test set a lot worse, should be explored
 - Note: explored time differencing but with little good results
- 3. Sentiment comments not included**
 - Focus on drivers the company can directly influence (but found to be relevant in literature)



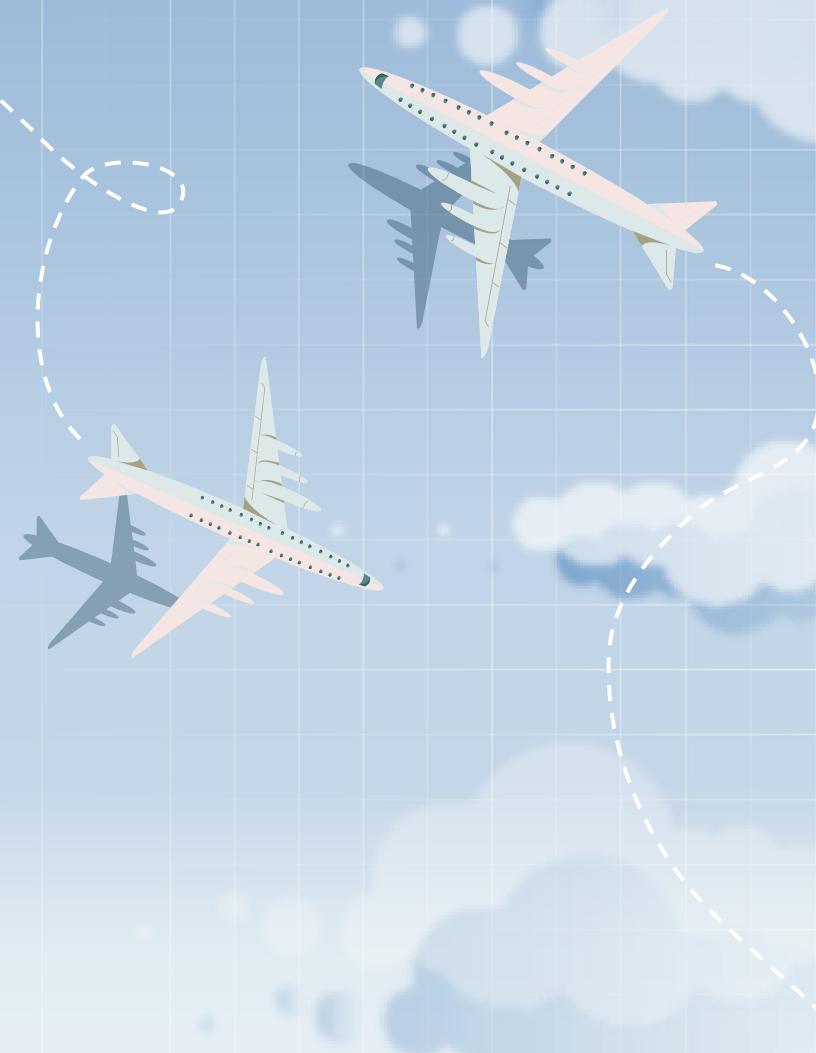
THANK YOU FOR YOUR ATTENTION

DO YOU HAVE ANY QUESTIONS?

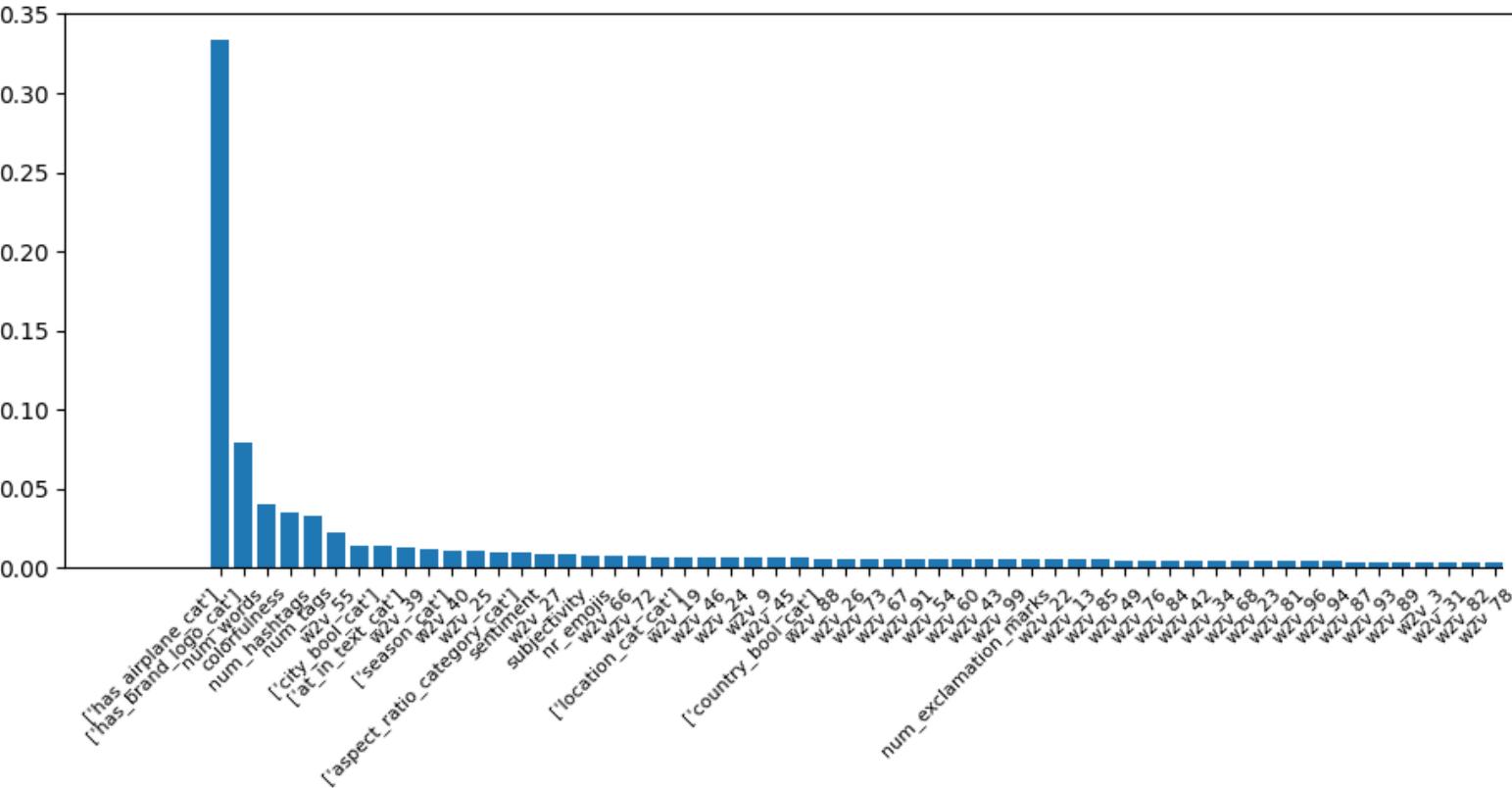
Ballegeer Matteo	01900129
De Rocker Yaël	01904043
Janssens Wannes	01905583
Peire Julian	01900199
Willemen Tom	01900194



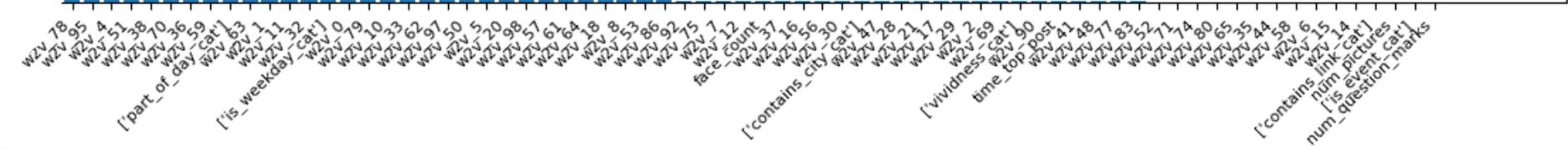
SUPPORTING SLIDES



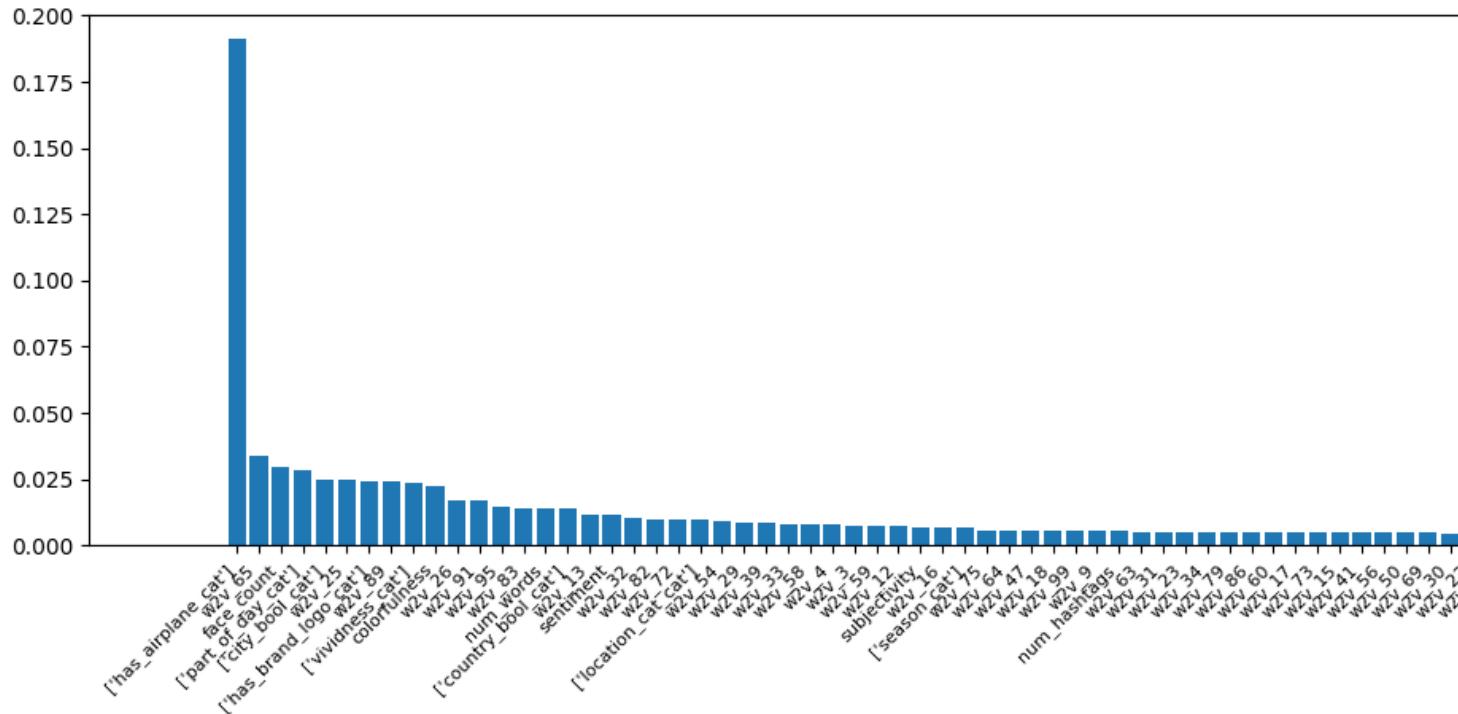
FEATURE IMPORTANCE LUFTHANSA USA



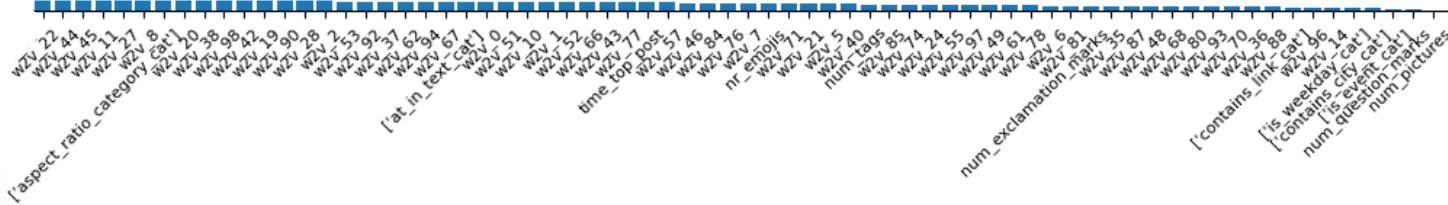
FEATURE IMPORTANCE LUFTHANSA USA (CONTINUED)



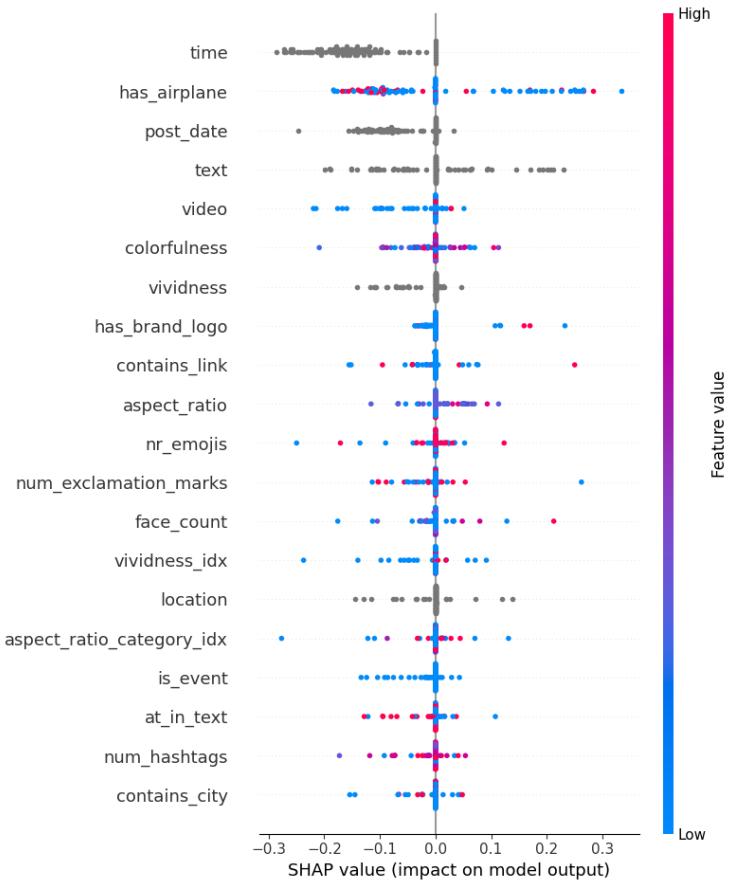
FEATURE IMPORTANCE SINGAPORE AIR



FEATURE IMPORTANCE SINGAPORE AIR (CONTINUED)



SHAPLEY SINGAPORE AIR



ENGAGEMENT VALENCE

