



Success Factors of Social Media Content – The Case Of Accenture Facebook Posts

RPV: Data Science Methods and Technologies
for Data-driven Business Models

St. Gallen, 21.12.2016

Valentin Kahn, 12-613-378

A solid blue vertical bar located to the left of the "Agenda" header.

Agenda

01

Method

02

Results

03

Conclusions
& Challenges

04

Discussion



Universität St. Gallen

01

Method

- Research Question
- Dataset & Subsets
- Research Process
- Key Queries



Research Question & Hypotheses

What do Accenture posts that create the most attention, the biggest discussion and the best emotions have in common?



User Engagement

H2
Questions
& calls to
action

H6
Media
elements

H4
Links

Emotions & Vibe

H5
Length &
time focus

H1
Brand-
related
content

H3
Questions

Dataset & Subsets

Correlation Analysis I

Analysis of all posts and comments (11'352 rows)



Analysis and correlation between:

- Length of posts
- Time tense of posts
- Links
- Questions or calls to action
- Brand-related content

and

- Likes
- Shares

Correlation Analysis II

Analysis of posts that have comments (2'140 rows)



Analysis and correlation between:

- Length of posts
- Time tense of posts
- Links
- Questions or calls to action
- Brand-related content

and

- Likes
- Shares
- Comments
- Emotions of comments

Best-in-Class Analysis

Analysis of «best-in-class» posts (and their comments; 168 rows)



Analysis and identification of success factors based on commonalities in:

- Content (words) of posts
- Length of posts
- Time tense of posts
- Questions and call to actions
- Brand-related content
- Links

Research Process & Key Queries

1

Exploring Dataset

2

Literature Review

3

Analysis

4

Triangulation

```
SELECT status_message, length(status_message), focuspast, focuspresent,
focusfuture, num_likes, num_shares, t.pos, t.neg
FROM accenture
INNER JOIN
(SELECT avg(posemo) as pos, avg(negemo) as neg, original_post_id
FROM accenture
WHERE type = "comment"
GROUP BY original_post_id)
as t
ON accenture.post_id_summary = t.original_post_id;
```

```
select avg(num_shares), avg(num_likes) from accenture where type = "post"
and num_likes < 1396069215
and
(status_message like "%X%" or
status_message like "%Accenture%" or
status_message like "%IT%" or
status_message like "%business%" or
status_message like "%companies%" or
status_message like "%digital%" or
status_message like "%technology%" or
status_message like "%services%" or
status_message like "%accenture's%")
union
select avg(num_shares), avg(num_likes) from accenture
where type = "post"
and num_likes < 1396069215
and (status_message not like "%Accenture%"
and status_message not like "%IT%"
and status_message not like "%business%"
and status_message not like "%companies%"
and status_message not like "%digital%"
and status_message not like "%technology%"
and status_message not like "%services%"
and status_message not like "%accenture's%")
;
```

```
SELECT SUM(total_count) as total, value
FROM
(SELECT count(*) AS total_count, REPLACE(REPLACE(REPLACE(x.value,'?', ''), ',', ''), '!', '') as value
FROM
(SELECT SUBSTRING_INDEX(SUBSTRING_INDEX(t.status_message, ' ', n,n), ' ', -1) value
FROM
(SELECT status_message FROM accenture
WHERE type = "post"
and num_likes > 5
and num_comments > 0
and num_shares > 0
and posemo > 5
and negemo = 0
and post_id IN
(SELECT Distinct original_post_id from accenture
where posemo > 5
and negemo = 0
and type = "comment"))) t
CROSS JOIN
(SELECT a.N + b.N * 10 + 1 n
FROM
(SELECT 0 AS N UNION ALL SELECT 1 UNION ALL SELECT 2 UNION ALL SELECT 3 UNION ALL SELECT 4
UNION ALL SELECT 5 UNION ALL SELECT 6 UNION ALL SELECT 7 UNION ALL SELECT 8 UNION ALL SELECT 9) a
,(SELECT 0 AS N UNION ALL SELECT 1 UNION ALL SELECT 2 UNION ALL SELECT 3 UNION ALL SELECT 4
UNION ALL SELECT 5 UNION ALL SELECT 6 UNION ALL SELECT 7 UNION ALL SELECT 8 UNION ALL SELECT 9) b
ORDER BY n) n
WHERE n.n <= 1 + (LENGTH(t.status_message) - LENGTH(REPLACE(t.status_message, ' ', '')))
ORDER BY x.value) AS x
GROUP BY x.value
ORDER BY total
DESC LIMIT 50;
```

02

Results



Research Results (1/2)

Content

Top Key Words	Occurrence	Top 168
1. Accenture	3'260	62
2. IT	1'347	
3. Business	993	21
4. Companies	865	23
....		

With Key Words

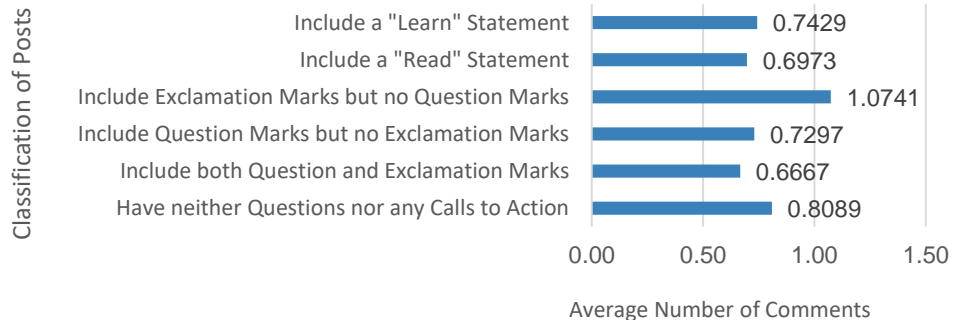
- 32.13 likes ($\sigma = 114.36$)
- 2.768 shares ($\sigma = 25.39$)

Without Key Words

- 36.32 likes ($\sigma = 117.07$)
- 2.717 shares ($\sigma = 11.01$)

Questions & Calls to Action

Average Number of Comments of Posts that...



With Question Mark

24 likes ($\sigma = 24.45$)

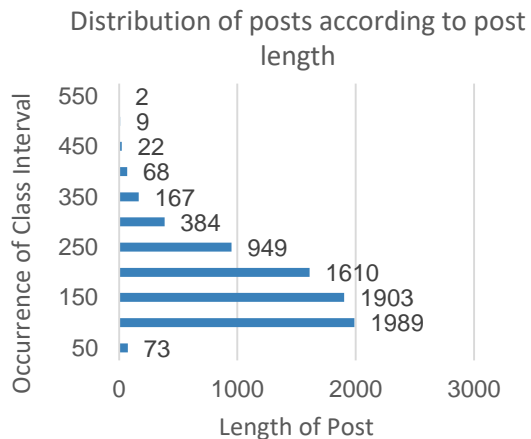
Without Question Mark

35 likes ($\sigma = 130.45$)

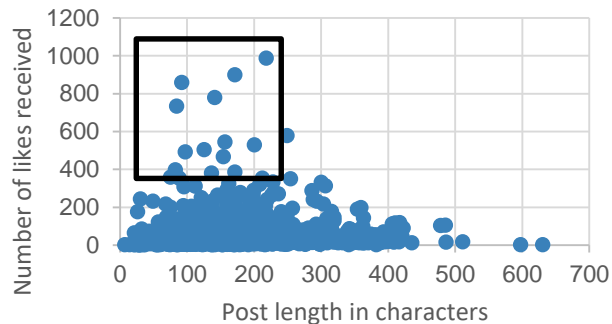
Research Results (2/2)

Links, Length & Time Focus

- No clear correlation between these factors and comments, shares or likes
- Optimal post length seems to be 80 to 250 characters



Optimal post length in terms of numbers of likes



Media Elements

- Clear correlation between media elements and number of likes, especially photos
- Same applies for comments
- From the 25 posts that received more than 400 likes, 16 were a photo

Element	Likes	Comments
Links	28	0.72
Photos	127	1.60
Videos	55	1.27

03

Conclusions & Challenges



Conclusions & Challenges

Conclusions

- Brand-related key words could not trigger success
- Calls to action: Only exclamation marks have a clear positive (but unclear) effect
- Links, post length and time focus do not trigger success either
- Media elements (videos and especially photos) are successful



Challenges

- Setup of virtual machine
- Speed of Hadoop
- Iterative process of building and abandoning hypotheses
- Word count query
- From key words to conclusions



Discussion



Thank you for your attention!

St. Gallen, 21.12.2016

Valentin Kahn, 12-613-378

Annex

Further Selected Key Queries

```
SELECT status_message as Content FROM accenture
WHERE type = "post"
and num_likes > 5
and num_comments > 0
and num_shares > 0
and posemo > 5
and negemo = 0
and post_id
IN
(SELECT Distinct original_post_id from accenture
where posemo > 5
and negemo = 0
and type = "comment");
```

Selection of "best-in-class" posts

```
SELECT SUM(total_count) as total, value
FROM
  (SELECT count(*) AS total_count, REPLACE(REPLACE(REPLACE(x.value, '?', ''), '.', ''), '!', ',') as value
  FROM
    (SELECT SUBSTRING_INDEX(SUBSTRING_INDEX(t.status_message, ' ', n.n), ' ', -1) value
    FROM
      (SELECT status_message FROM accenture
      WHERE type = "post"
      and num_likes > 5
      and num_comments > 0
      and num_shares > 0
      and posemo > 5
      and negemo = 0
      and post_id IN
        (SELECT Distinct original_post_id from accenture
        where posemo > 5
        and negemo = 0
        and type = "comment")) t
    CROSS JOIN
      (SELECT a.N + b.N * 10 + 1 n
      FROM
        (SELECT 0 AS N UNION ALL SELECT 1 UNION ALL SELECT 2 UNION ALL SELECT 3 UNION ALL SELECT 4
        UNION ALL SELECT 5 UNION ALL SELECT 6 UNION ALL SELECT 7 UNION ALL SELECT 8 UNION ALL SELECT 9) a
        ,(SELECT 0 AS N UNION ALL SELECT 1 UNION ALL SELECT 2 UNION ALL SELECT 3 UNION ALL SELECT 4
        UNION ALL SELECT 5 UNION ALL SELECT 6 UNION ALL SELECT 7 UNION ALL SELECT 8 UNION ALL SELECT 9) b
        ORDER BY n) n
      WHERE n.n <= 1 + (LENGTH(t.status_message) - LENGTH(REPLACE(t.status_message, ' ', '')))
      ORDER BY value) AS x
    ORDER BY x.value) AS y
  GROUP BY value
  ORDER BY total
  DESC LIMIT 50;
```

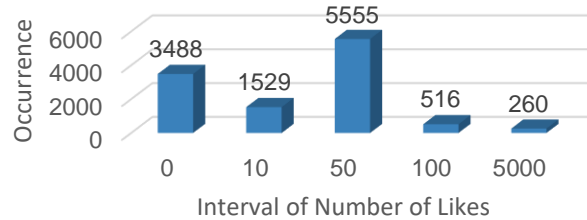
word count (top 50) in "best-in-class" posts

```
select avg(num_comments) from accenture where type = "post" and status_message not like "%!" and status_message not like "%?%"
and status_message not like "%Read%" and status_message not like "%Learn%"
union
select avg(num_comments) from accenture where type = "post" and status_message like "%!" and status_message like "%?%"
union
select avg(num_comments) from accenture where type = "post" and status_message like "%?%" and status_message not like "%!"
union
select avg(num_comments) from accenture where type = "post" and status_message like "%!" and status_message not like "%?%"
union
select avg(num_comments) from accenture where type = "post" and status_message like "%Read%"
union
select avg(num_comments) from accenture where type = "post" and status_message like "%Learn%";
```

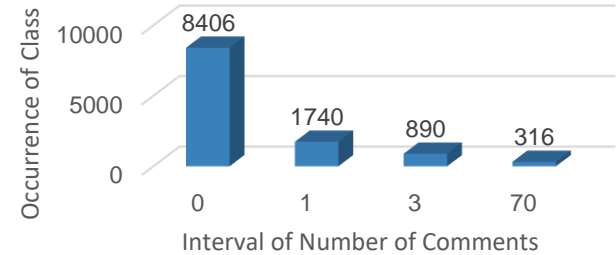
exploring the relation between questions, calls to actions and user engagement

Further Statistics

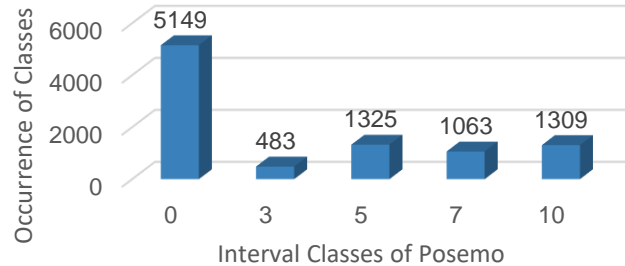
Distribution of Number of Likes per Post



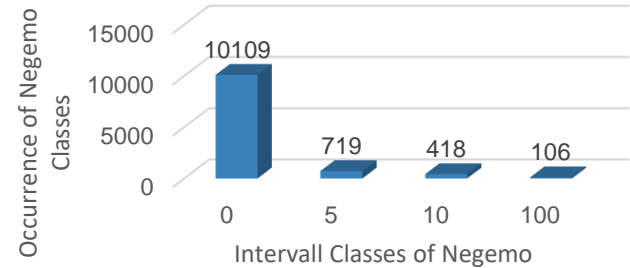
Distribution of Number of Comments



Distribution of Posemo Values



Distribution of Negemo Values





Universität St. Gallen

accenture

Success Factors of Accenture Facebook Posts

RPV: Data Science Methods and Technologies
for Data-driven Business Models

St. Gallen, 21.12.2016

Valentin Kahn, 12-613-378