

Osumi dojo report -2015-

Iguchi Nagahama Kawaguchi
Pratidina Yuan

Study about Cheese

Our point of interests:

- “How is Cheese being viewed by the Japanese?”
- “Will the conversation in the Net (ネット) influence the consumption of cheese in Japan?”
 - correlation between ネット and consumption
 - comparison with other product (butter)
- “Cheese Festival”
- Area & Cheese

Method

口コミ件数を求め、推移を消費量と比較する

Source of Data:

1. Kuchikomi@kakaricyo (口コミ)

→ as sample for conversation in the Net

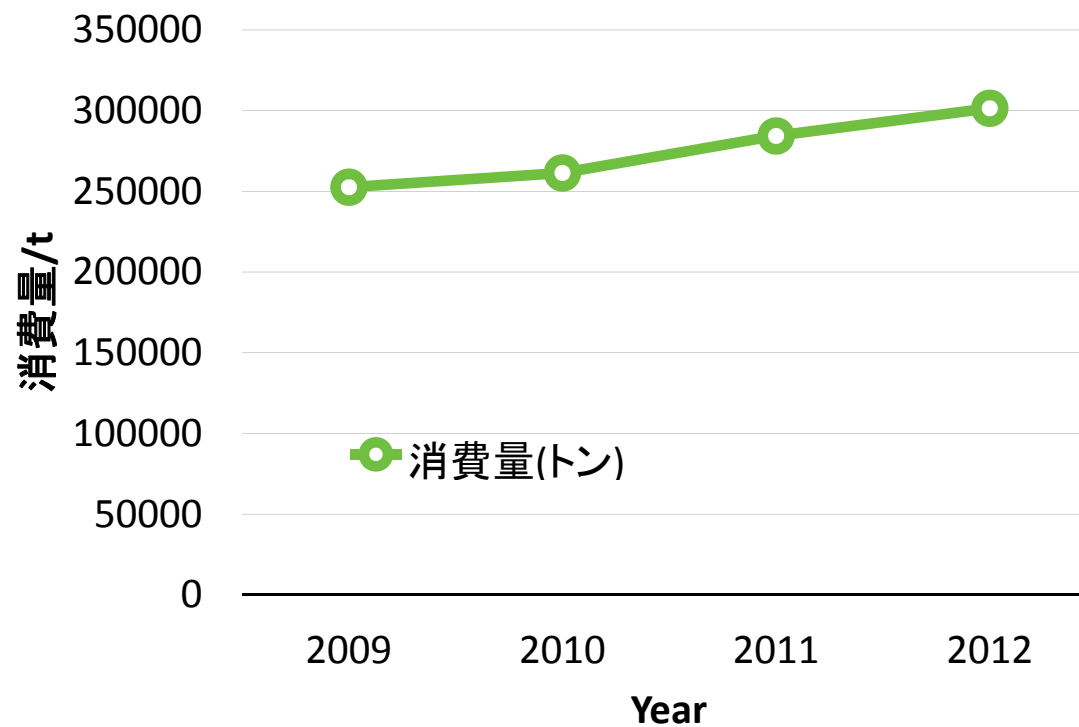
2. 外部データ（農林生産省の消費量、生産量）

口コミ係長

- Use d チーズ as keyword to search in "Blog, 2Ch, 掲示板"
- Set the time-frame per 3 months in period of 2009-2012 then accumulated the number of hits
- Set the SPAM filter into スпам強+広告系

1. Cheese Consumption (2009-2012)

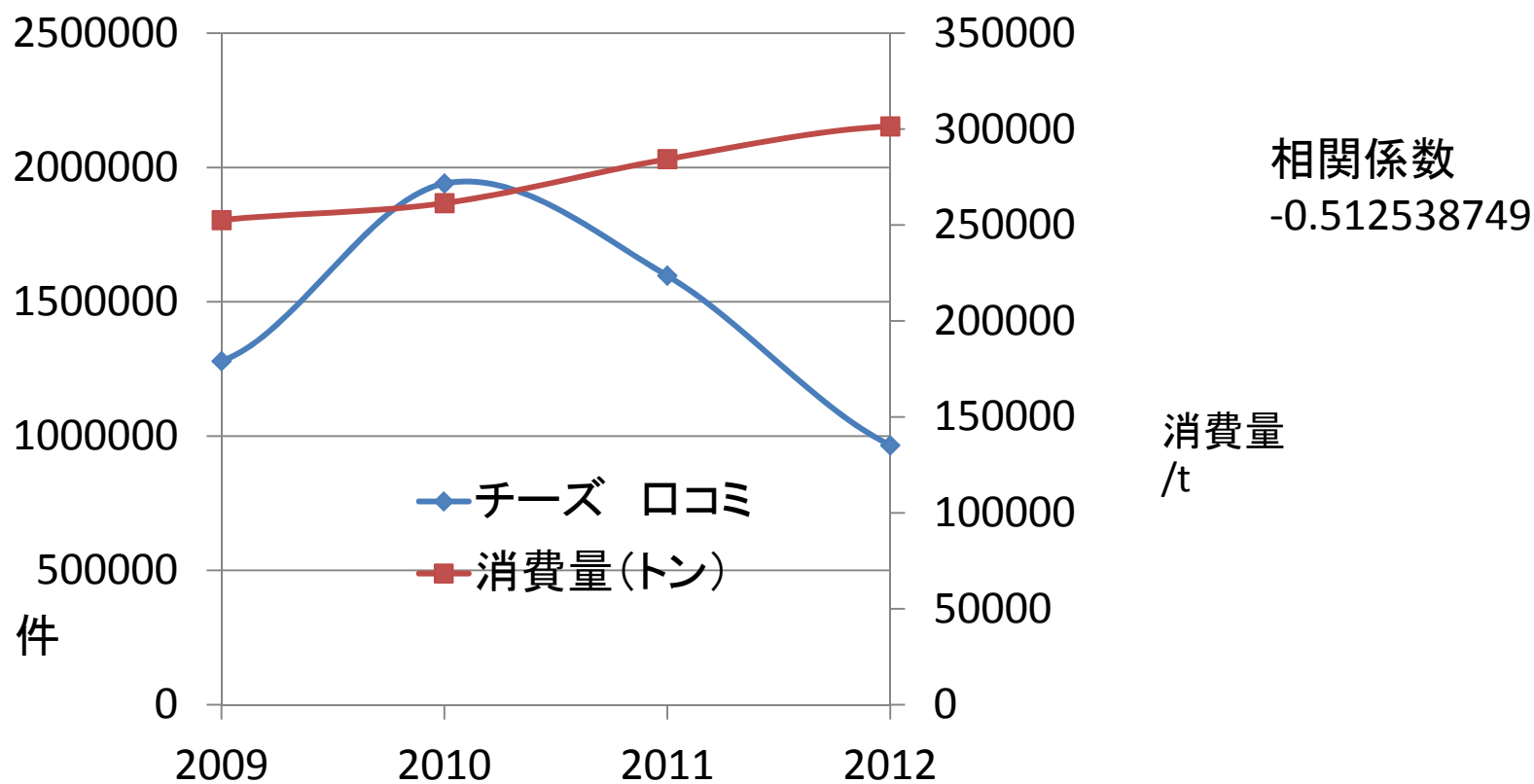
- 増加傾向がある



Source: 2012 Data

2. Cheese Consumption (2009-2012)

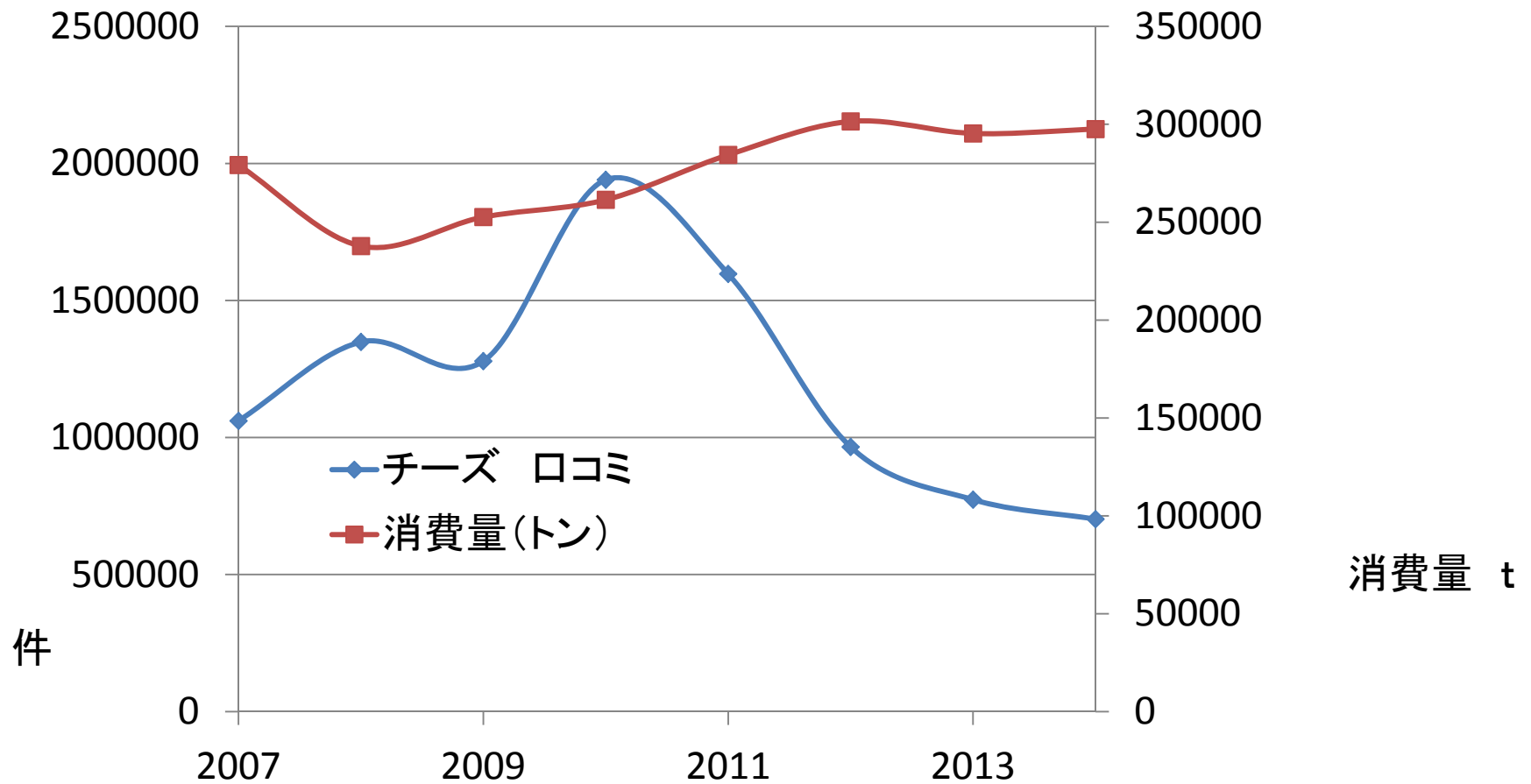
- 各年の口コミ数との相関はあまりなかった



3. Cheese Consumption (2007-2014)

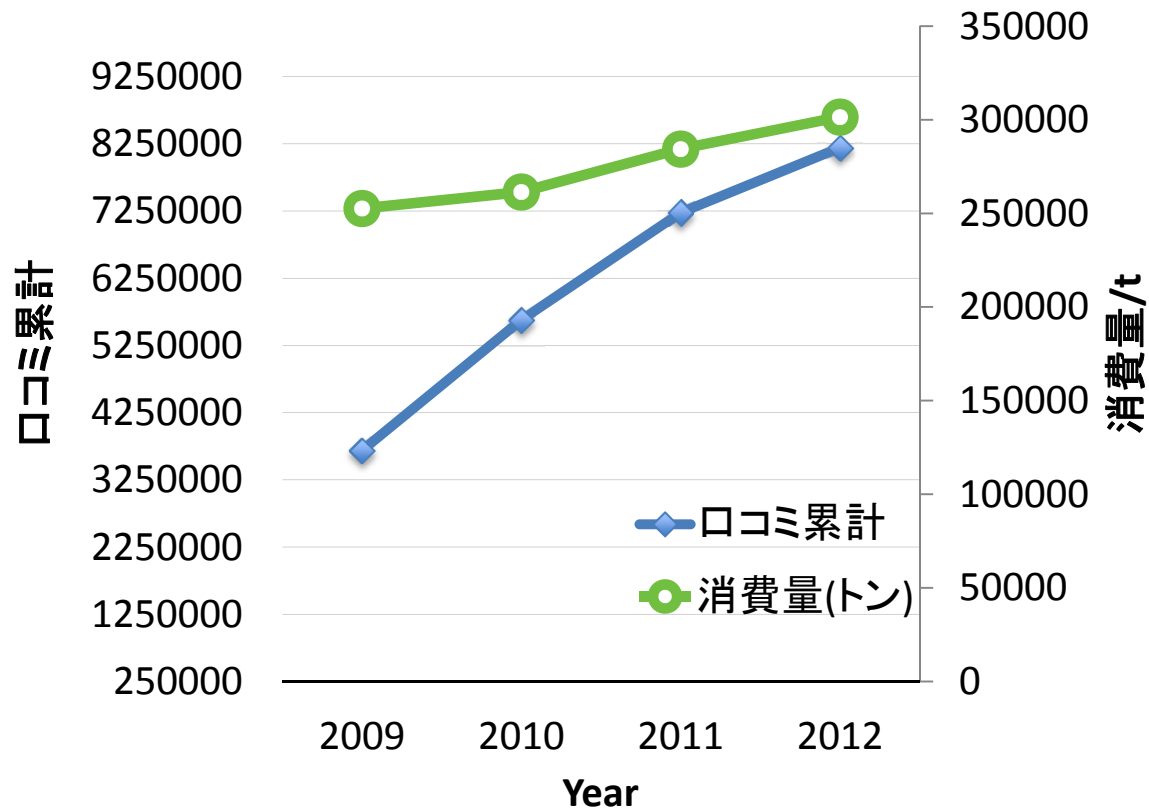
相関係数
-0.590192566

- 期間延ばして相関はあまりなかった



4. Cheese Consumption & Kuchikomi (2009-2012)

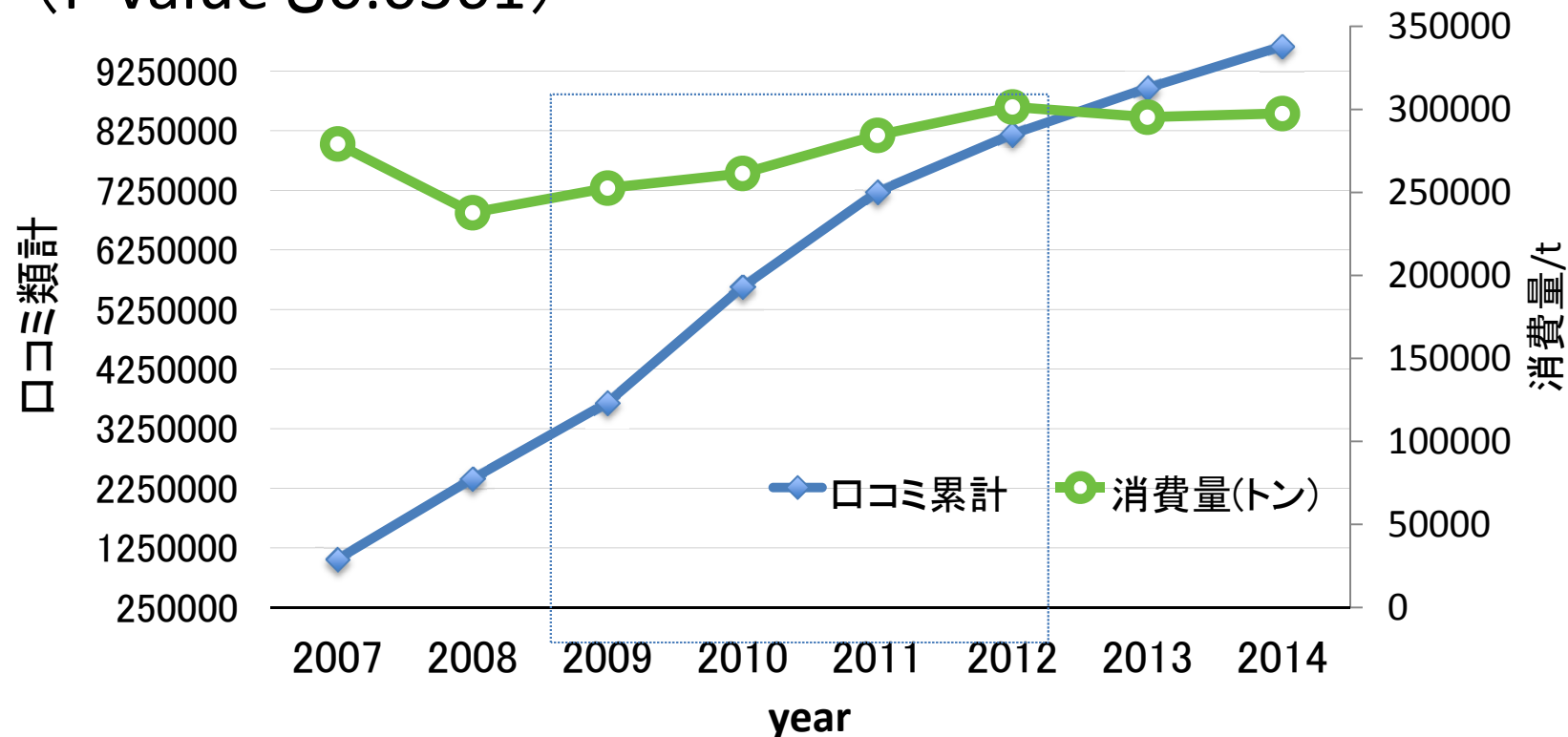
- 口コミ累計と消費量の相関係数0.963



→期間を延ばしてみた

5. Cheese Consumption & Kuchikomi (2007-2014)

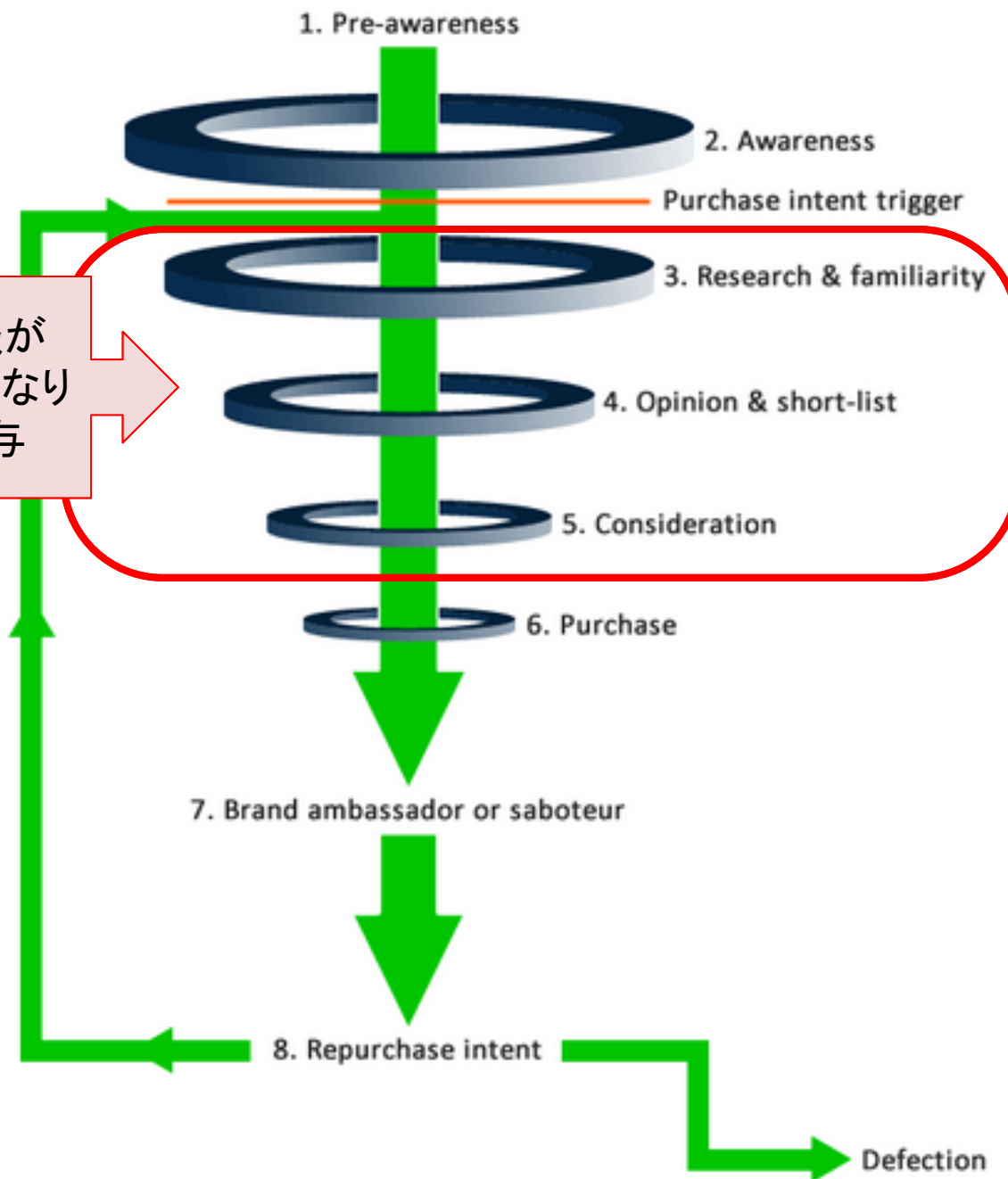
- 口コミ累計と消費量の相関係数0.739
(期間延ばしてもそれなりに相関がある)
(P valueも0.0361)



Discussion

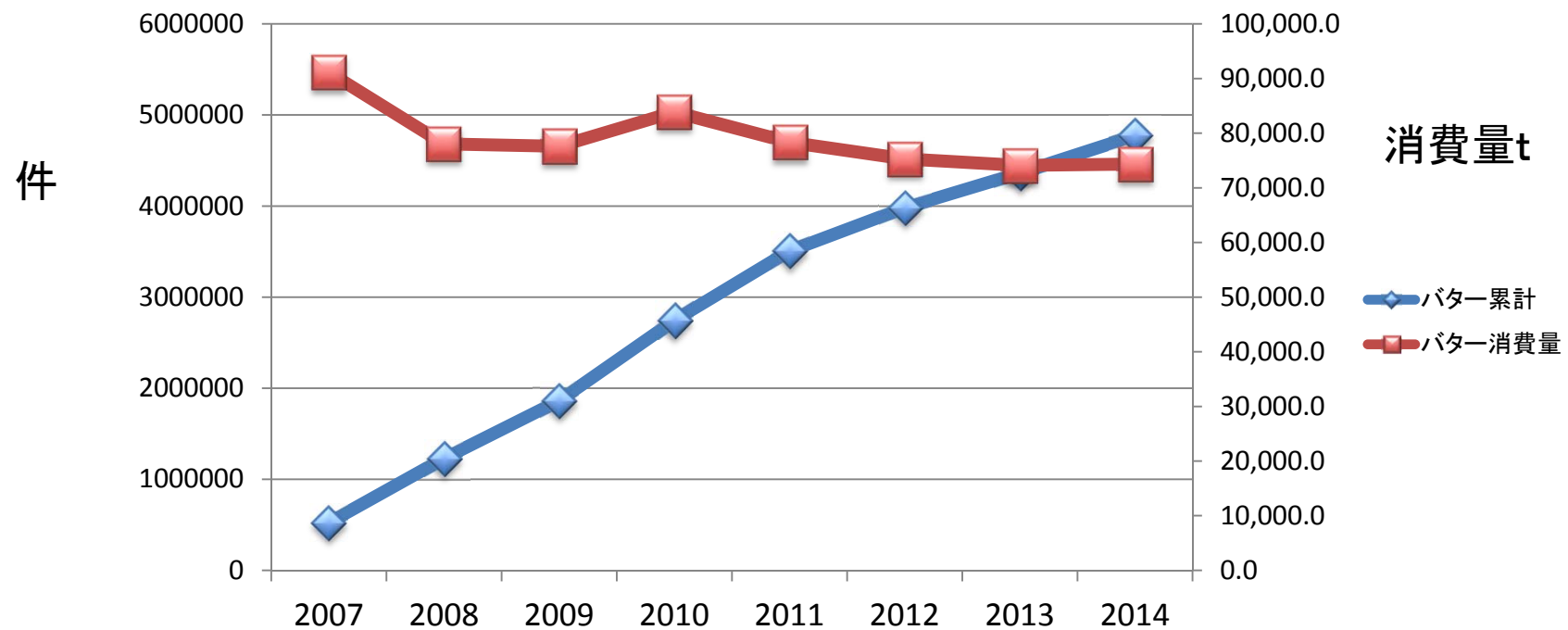
- 口コミは累積で消費量に影響する
- 消費者の傾向 (Purchase funnel) の中で寄与していると考えられる部分

口コミ情報が
長年積み重なり
ここに寄与



OTHER data

- 他の食品についても同様にデータの収集、解析
 - 口コミ累計と消費量の相関係数 -0.749



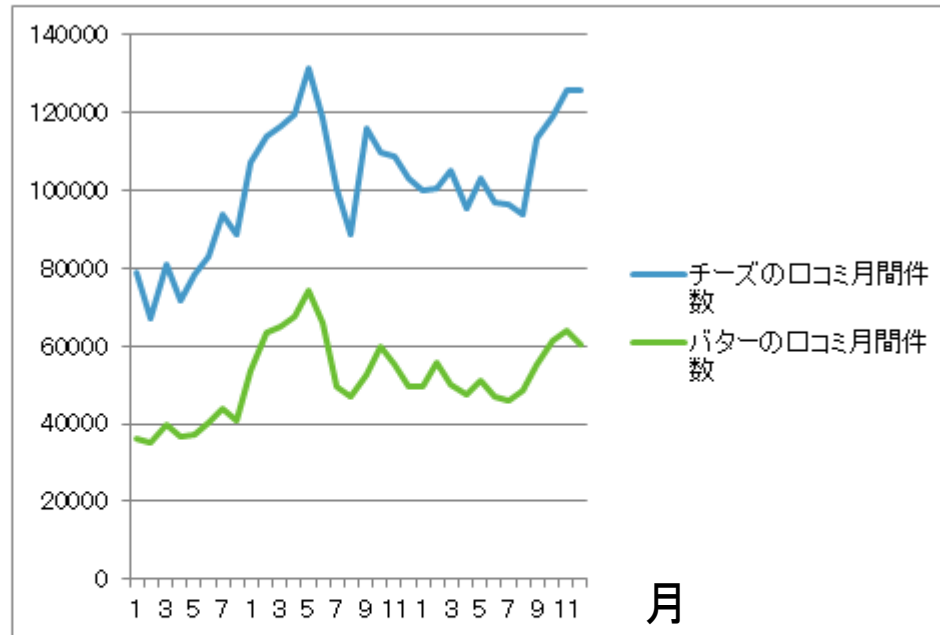
データソース: 口コミ@係長、一般社団法人「ミルク」主要国の消費動向」

チーズとバターのココミ月間件数の回帰分析

概要

回帰統計	
重相関 R	0.9301
重決定	0.865
補正 R2	0.8609
標準誤差	5692.1
観測数	35

件



月

分散分析

	自由度	変動	分散	観測され	有意 F
回帰	1	7E+09	7E+09	211.43	7E-16
残差	33	1E+09	3E+07		
合計	34	8E+09			

	係数	標準誤差	t	P-値	下限	上限	下限	上限	
切片	24309	5461.3	4.4511	9E-05	13198	35420	13198	35420	
	36093	1.4965	0.1029	14.541	7E-16	1.2871	1.7059	1.2871	1.7059

同じ乳製品としてココミトレンドは似ているもの、消費との相関は異なる
 →消費にはココミ以外の要因もあることが示された

OTHER data

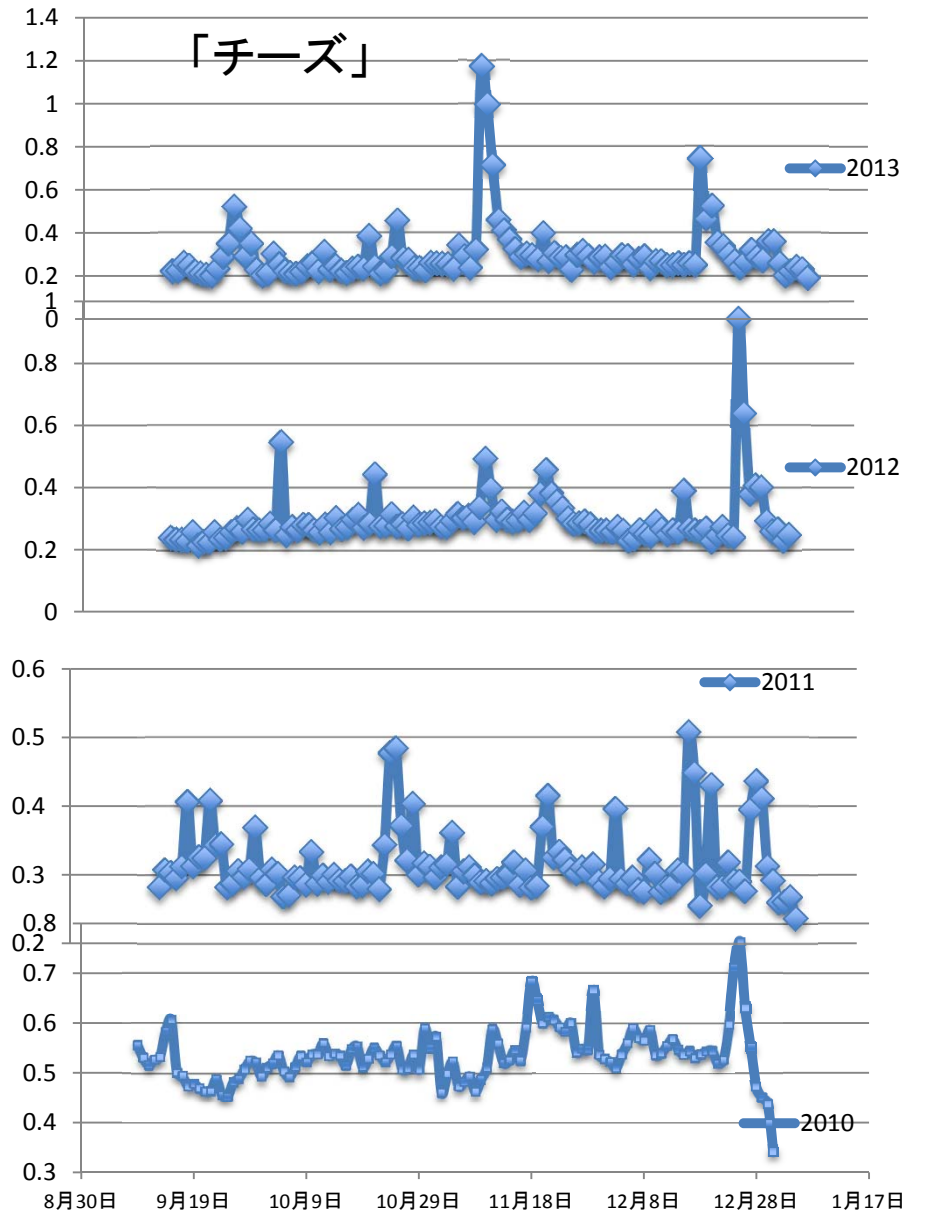
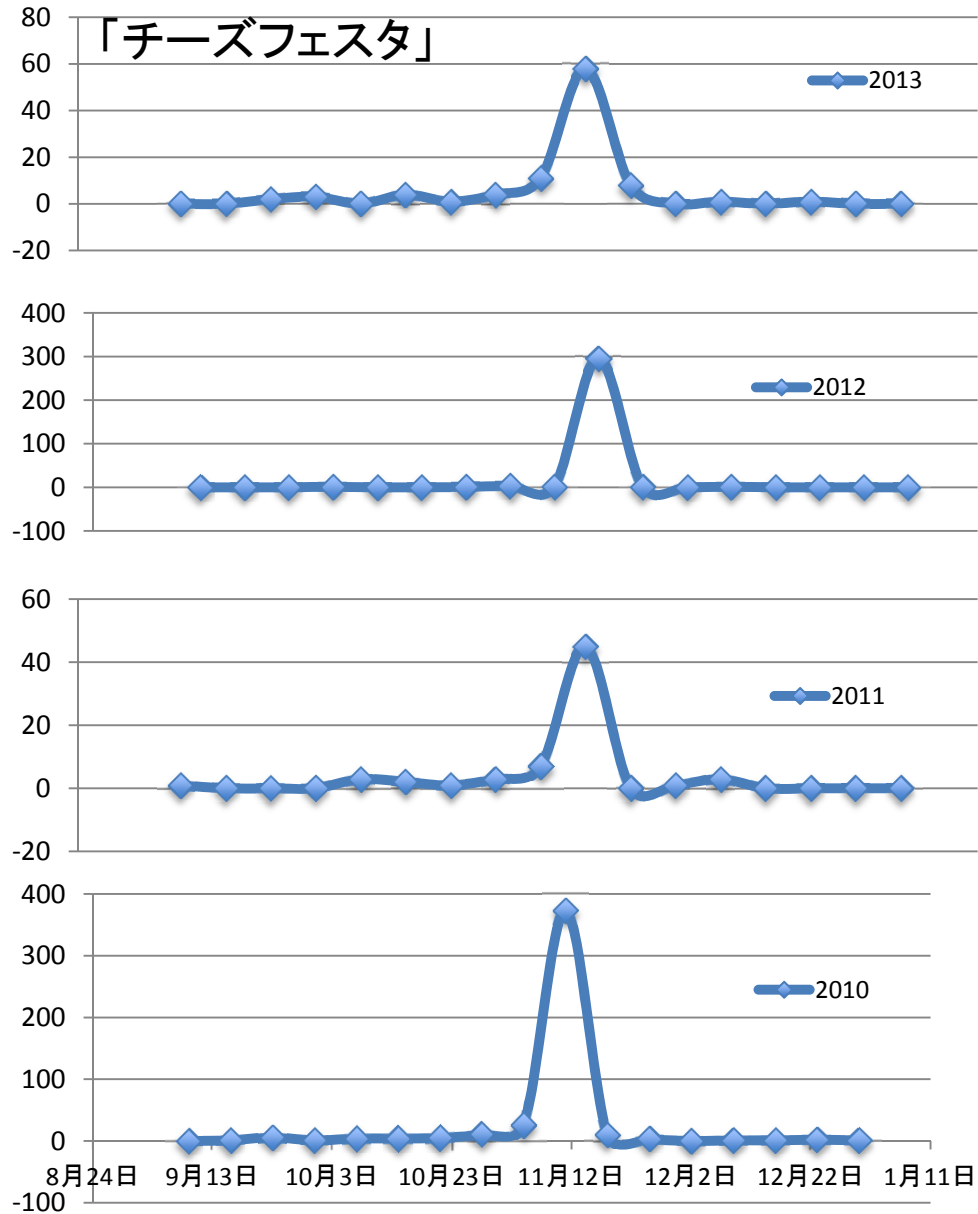
- Area別の消費量vs
ロコミ数
→ロコミ数全般(、)の
地域性の影響が大きい
(相関係数が約1)

2012	ロコミ数	
検索条件	チーズ	、
北海道	1846	147331
東北	1502	134378
関東	15102	1334484
中部	4430	413901
近畿	6291	527534
中国	1355	127593
四国	723	64165
九州沖縄	2481	232850
海外その他	4323	381713
	相関係数	0.9992786

	Spring	Summer	Autumn	Winter
2013	0.999279	0.999502	0.999171	0.999642
2014	0.999391	0.999323	0.998592	0.998906

→ロコミ係長でチーズのロコミの地域性把握は
困難

OTHER data



Conclusion

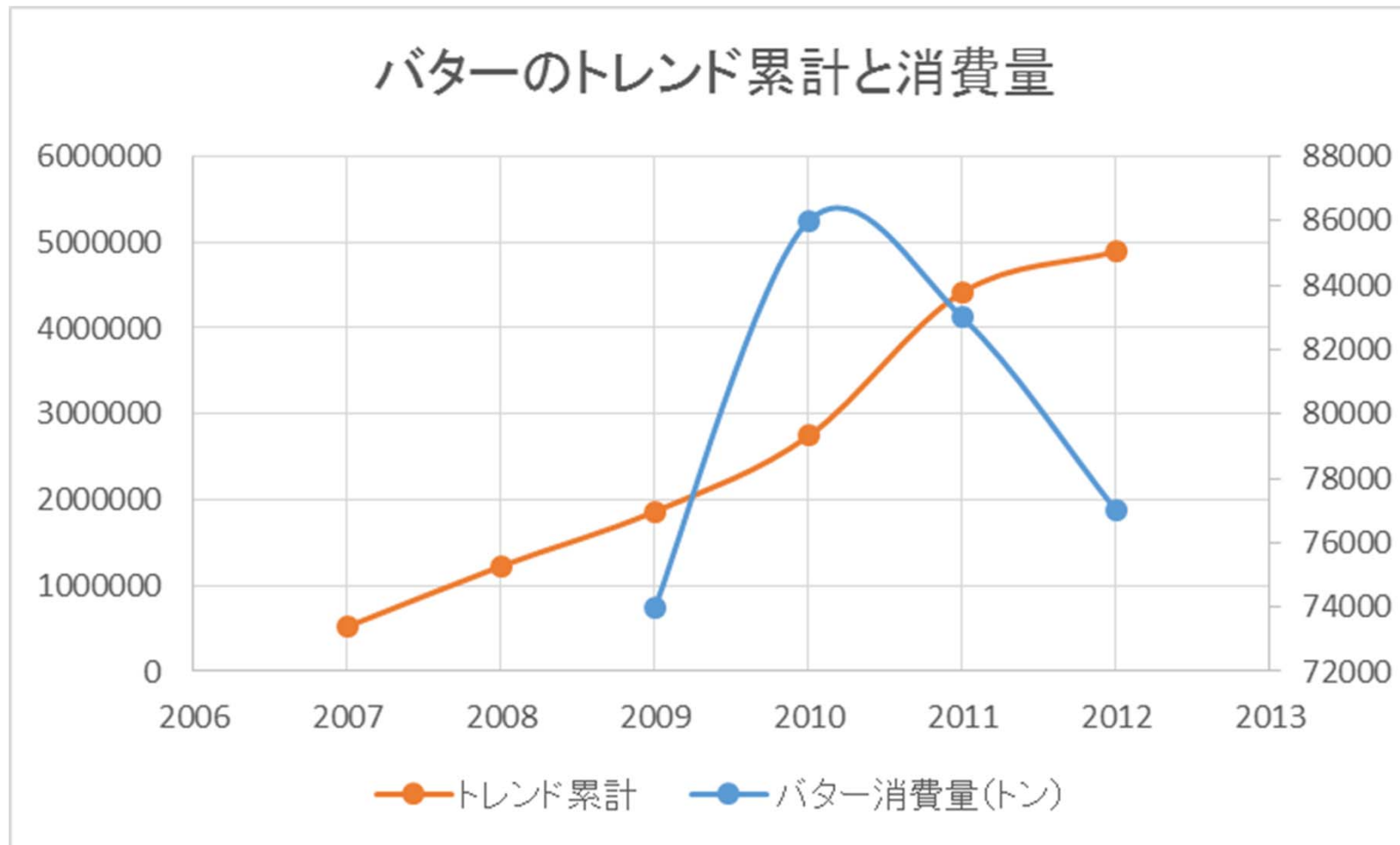
- 累計が消費量と相関があるのは当たり前？
(両方とも増加だから→消費が減っているものについて今後要検討)
- ロコミ係長でチーズのトレンドの地域性を把握するのは難しい
- 同じ乳製品でも消費量とロコミ数の相関は異なる
- チーズフェスタのロコミへの影響は少ない

感想

- チーズについて何を調べるか、topicを決めるのが難しかった
- 思ったより口コミを拾いづらく、データ収集に苦労した
- データが集まってもどのような切り口で考察すればいいのかが難しかった
- Big data is challenging!

OTHER data

- 他の食品についても同様にデータの収集、解析



データソース:ロコミ@係長、一般社会法人「ミルク」主要国の消費動向」