

ビッグデータを活用した市民意識の調査・分析業務
報告書

令和2年3月

株式会社ノーザンクロス

目 次

第1章 調査・分析の概要

- 1. 調査・分析の目的..... I - 1
- 2. 実施概要..... I - 1

第2章 ソーシャルメディア等における市民等のニーズ・意見等の抽出及び分析

抽出・分析の方法について

- 1. 概要..... II - 1
- 2. Twitter 分析について..... II - 1
- 3. コールセンターへの問合せの分析について II - 5
- 4. 分析に用いた図の説明と凡例 II - 7
- 5. 分析についての注記 II - 9

札幌の全体イメージ について

- 1. Twitter データ II - 10
- 2. 分析..... II - 17

分野1 地域～地域での支え合いとつながりづくり について

- 1. Twitter データ II - 20
- 2. コールセンターへの問合せ II - 25
- 3. 分析..... II - 29

分野2 経済～暮らしと雇用を支える経済の発展 について

- 1. Twitter データ II - 32
- 2. コールセンターへの問合せ II - 38
- 3. 分析..... II - 42

分野3 子ども・若者～将来を担う子ども・若者の健やかな育み について

- 1. Twitter データ II - 46
- 2. コールセンターへの問合せ II - 52
- 3. 分析..... II - 56

分野4 安全・安心～安心して暮らせる「人に優しい」まちづくり について

- 1. Twitter データ II - 59
- 2. コールセンターへの問合せ II - 64
- 3. 分析..... II - 68

分野5 環境～次世代へつなげる持続可能なまちづくり について

- 1. Twitter データ II - 72

2. コールセンターへの問合せ	II - 77
3. 分析	II - 81

分野6 文化～文化芸術・スポーツによる創造性の育み について

1. Twitter データ	II - 86
2. コールセンターへの問合せ	II - 92
3. 分析	II - 96

分野7 都市空間～魅力と活力を持続的に高める集約型のまちづくり について

1. Twitter データ	II -101
2. コールセンターへの問合せ	II -106
3. 分析	II -110

第3章 今後のまちづくりにおける重要な視点・取組・分野の提案～

分析を踏まえた気づきと、基本目標の提案

札幌の全体イメージ について.....	III - 1
分野1 地域～地域での支え合いとつながりづくり について	III - 1
分野2 経済～暮らしと雇用を支える経済の発展 について.....	III - 3
分野3 子ども・若者～将来を担う子ども・若者の健やかな育み について.....	III - 5
分野4 安全・安心～安心して暮らせる「人に優しい」まちづくり について.....	III - 7
分野5 環境～次世代へつなげる持続可能なまちづくり について	III - 9
分野6 文化～文化芸術・スポーツによる創造性の育み について	III - 11
分野7 都市空間～魅力と活力を持続的に高める集約型のまちづくり について	III - 13
分野横断的な視点.....	III - 15

第4章 アンケート調査の企画と調査票案の設計

1. アンケート調査の企画.....	IV - 1
2. アンケート調査票案	IV - 2

第5章 ニーズ・意見等の分析・整理結果を行政計画の策定時に活用する手法の提案

1. Twitter データ、札幌市コールセンターデータの特徴.....	V - 1
2. 行政計画策定等への活用の方向性	V - 2
3. 行政計画策定におけるフレーム（様式）の提案	V - 3

第1章 調査・分析の概要

1. 調査・分析の目的

札幌市では「札幌市まちづくり戦略ビジョン」（計画期間：平成25～34年度。以下、戦略ビジョン）の推進にあたって「市民が主役のまちづくり」を基本姿勢の1つに掲げ、市民自治のまちづくりを進めることとしており、市民感覚を大切に、市民力を結集した行政運営を進めるには、市民のニーズや意見を的確に把握し効果的・効率的に政策に反映させることが求められる。

そのため、少子高齢化の進展や情報発信手段の多様化など、市民を取り巻く環境に大きな変化が生じている状況の中、ソーシャルメディア上の情報などのビッグデータから市民等が自発的に発信した意見を収集・分析することで、潜在的なニーズの発掘につなげ、今後の先導的な施策の展開につなげていくことを目的とし、本事業を実施した。

2. 実施内容

（1）ソーシャルメディア等における市民等のニーズ・意見等の抽出及び分析

2019年にソーシャルメディア「Twitter」において「札幌」に関して投稿されたコメント（以下、ツイート）を抽出し分析した。また、2018年度における札幌市コールセンターに寄せられた問合せについて分析した。

（2）今後のまちづくりにおける重要な視点・取組・分野の提案

（1）の分析を踏まえ、今後の行政計画策定・施策検討等に関して得られた気づきと、現行の戦略ビジョンにおける「7つの分野」ごとに設定された「基本目標」の修正案、その内容（項目）の案を検討・提案した。

（3）今後の行政計画の策定時におけるアンケート調査の企画と調査票案の設計

（1）（2）を踏まえ、今後の行政計画策定等に際して市民のニーズ・意見を調査するためのアンケート調査の企画と調査票案の設計を行った。

（4）ニーズ・意見等の分析・整理結果を行政計画の策定時に活用する手法の提案

（1）（3）で分析・整理した市民のニーズ・意見等を今後の行政計画等の策定時に活用する手法及びフレーム（様式）を検討・提案した。

第2章 ソーシャルメディア等における 市民等のニーズ・意見等の抽出及び分析

抽出・分析の方法について

1. 概要

市民等のニーズ・意見等を探るため、2019年のTwitterにおける「札幌」に関するツイート、2018年度における札幌市コールセンターに寄せられた問合せ（以下、コールセンターへの問合せ）について分析した。

2. Twitter 分析について

(1) データの抽出方法

- ・2019年の日本語による全ツイートの10%データ（推計約50億件）から「札幌」を意味する下記語句を含むツイート810,463件を抽出した。

札幌 さっぽろ サッポロ SAPPORO (※半角・全角、大文字・小文字は同じ文字として扱われる)
--

- ・なお、抽出にあたっては、定型文や宣伝文を自動でツイートするbotや、店舗・企業による宣伝・販売目的のツイート、風俗・出会いを目的としたツイート等（以下、スパムツイート）を削除するためのフィルタ（除外設定）を適用した。
- ・上記から、Twitterデータ使用の契約上の制限により、ランダムに150,000件のツイートを抽出した。
- ・上記データには、未だスパムツイートが含まれていたため、さらなるフィルタを設け、これらのツイートの除外を試みた。結果、138,835件のツイートを抽出できた。

(2) 分析データの抽出方法

- ・札幌の全体イメージと、現行戦略ビジョンにおけるまちづくりの7つの分野「地域」「経済」「子ども・若者」「安全・安心」「環境」「文化」「都市空間」について、分析を行った。
- ・7分野については、上述した140,891件のツイート（「環境」分野については138,835件のツイート）から、現行の戦略ビジョンの内容を踏まえ、各分野について数十個の語句を抽出語として設定し、その語句を含むツイートを抽出した。

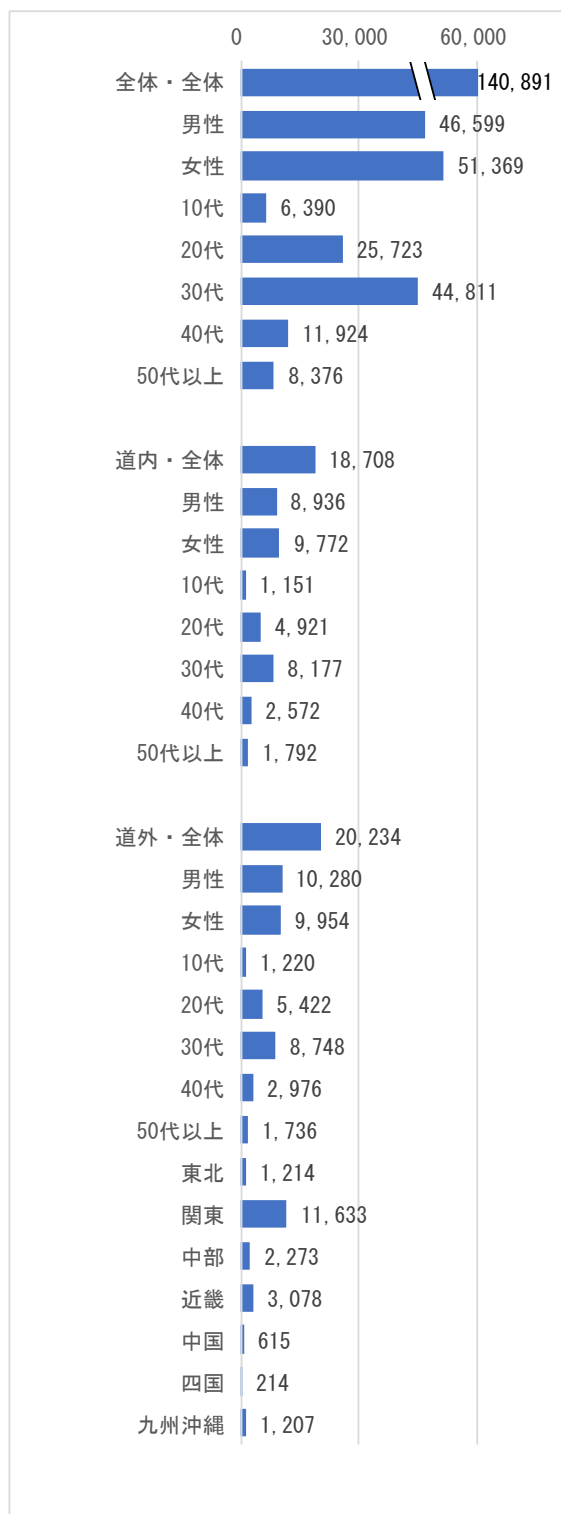
現行戦略ビジョンにおけるまちづくりの7つの分野

7つの分野	重要な視点	24の基本目標
地域	地域での支え合いとつながりづくり	<ul style="list-style-type: none"> ① 共生と交流により人と人がつながるまちにします ② 様々な担い手が地域のまちづくり活動に参加するまちにします ③ 多様な地域課題を解決できるまちにします
経済	暮らしと雇用を支える経済の発展	<ul style="list-style-type: none"> ④ 強みを生かした産業が経済をけん引するまちにします ⑤ 様々な連携により産業が高度化するまちにします ⑥ 市民の雇用が安定的に確保されるまちにします ⑦ 強みを生かし世界とつながるまちにします ⑧ 地域コミュニティを支える産業を大切にすまちにします
子ども・若者	将来を担う子ども・若者の健やかな育み	<ul style="list-style-type: none"> ⑨ 安心して子どもを生み育てられるまちにします ⑩ 将来を担う子どもの成長と自立を支えるまちにします ⑪ 若者が社会的に自立し活躍できるまちにします
安全・安心	安心して暮らせる「人に優しい」まちづくり	<ul style="list-style-type: none"> ⑫ 誰もが健康的で安心して暮らせるまちにします ⑬ 地域防災力が高く災害に強いまちにします ⑭ 安全な日常生活が送れるまちにします
環境	次世代へつなげる持続可能なまちづくり	<ul style="list-style-type: none"> ⑮ 豊かな自然と共生するまちにします ⑯ 資源やエネルギーを有効活用するまちにします ⑰ 市民が環境について学び行動するまちにします
文化	文化芸術・スポーツによる創造性の育み	<ul style="list-style-type: none"> ⑱ 創造的な活動により活力あふれるまちにします ⑲ 文化芸術やスポーツの魅力によりにぎわいが生まれるまちにします ⑳ 市民一人一人が魅力を再認識し発信するまちにします
都市空間	魅力と活力を持続的に高める集約型のまちづくり	<ul style="list-style-type: none"> ㉑ 公共交通を中心とした集約型のまちにします ㉒ 札幌の顔となる魅力と活力あふれる都心にします ㉓ 都市の価値を高めるみどりを生かしたまちにします ㉔ 都市基盤が適切に維持・保全されるまちにします

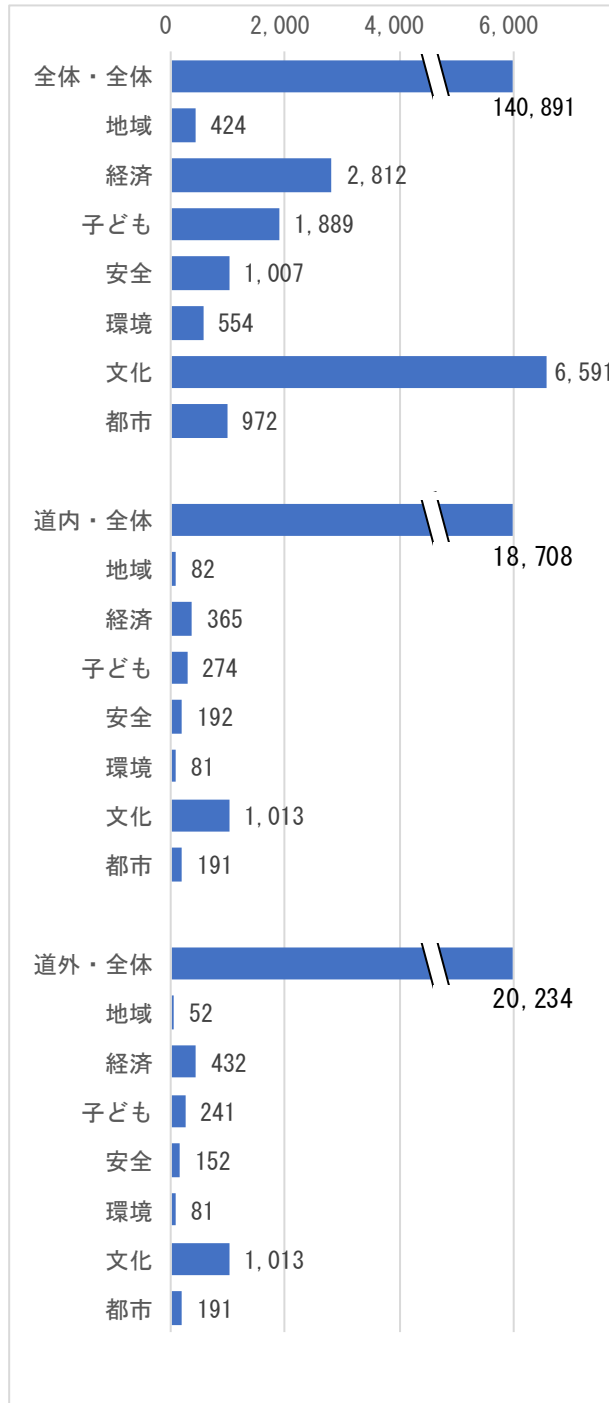
(3) プロファイル推定について

- Twitter データには性・年代・居住地等の属性（プロファイル）データは含まれていない。本分析においては、同一アカウントにおける過去のツイートの傾向からこれらの属性を推定する「プロファイル推定」を行った。これによって付与した属性により、データ全体に関する分析（以下、＜全体＞）に加え、道内居住者（以下、＜道内＞）、道外居住者（以下、＜道外＞）、別、性・年代・居住地域（国内8地方区分）別の分析を試みた。

プロファイル推定による属性別のツイート件数



<全体><道内><道外>別、
7つの分野について抽出したツイートの件数



3. コールセンターへの問合せの分析について

(1) 分析に用いたデータ

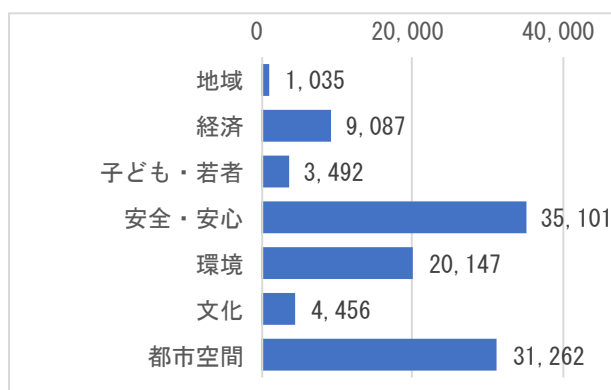
- ・元データは2018年度に札幌市コールセンターに寄せられた145,360件の問い合わせである。

(2) データの分類・分析方法

- ・元データから、無言電話、間違い電話、担当部署への転送依頼・連絡、イベントへの参加申込等を削除した。結果、データ件数は117,141件となった。
- ・元データは、内容を踏まえて、21の「カテゴリ」に分類されていた。本分析においては、これらの「カテゴリ」を下記のように戦略ビジョンの7つの分野に振り分けて分野ごとに分析を行った。
- ・なお、コールセンターへの問合せについては全体的な分析結果を掲載していないが、それは、全体的な分析結果が、分野ごとの分析の集約に過ぎなかったためである。
- ・また、コールセンターへの問い合わせの特色として、コールセンターが所管課の代行として問い合わせ及び申し込み受付を担っている、ハチの巣の駆除、高齢者肺炎球菌ワクチン接種の手続き等、コールセンターへの入電が特に多い事業がデータ件数に含まれていることを注記する。

カテゴリ→7つの分野への振り分け

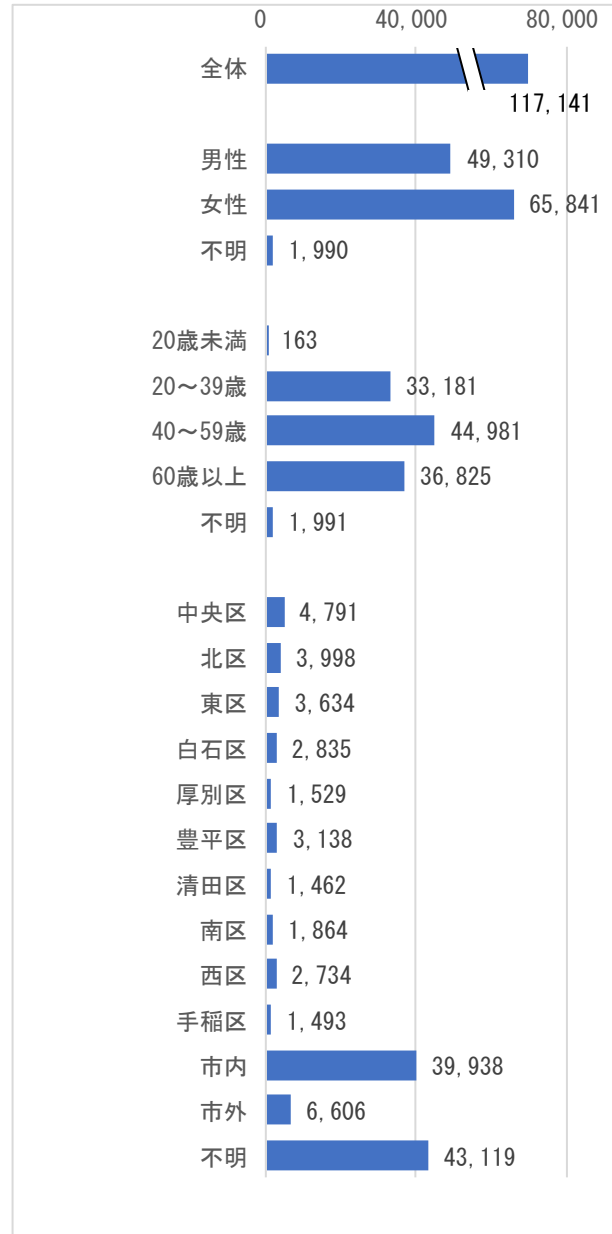
カテゴリ	件数	7分野への振り分け	件数
健康・医療・衛生	14,502	安全・安心	14,502
イベント・観光・スポーツ	4,456	文化	4,456
福祉・保険年金	17,123	安全・安心	17,123
公共交通・交通安全	8,318	都市空間	8,318
戸籍・証明	17,012	都市空間	17,012
ごみ・リサイクル・清掃	14,958	環境	14,958
教育・生涯学習	3,492	子ども・若者	3,492
除雪	2,124	安全・安心	2,124
公共施設	3,421	都市空間	3,421
税金	7,506	経済	7,506
動物（犬・猫・カラス・ハチ・熊等）	4,152	環境	4,152
防災対策・消防	1,352	安全・安心	1,352
法律相談・生活相談	1,581	経済	1,581
住宅	1,035	地域	1,035
自然（山・川）・公園	1,037	環境	1,037
道路・街路灯	1,610	都市空間	1,610
上・下水道	901	都市空間	901
English	406		
韓国語	69		
中国語	86		
その他	12,000		
計	117,141		



(3) 属性について

- ・コールセンターへの問合せデータには、性・年代・居住地（区）の属性が記録されていた。これらに基づいて、各属性別の分析を試みた。

属性別の問合せ件数



4. 分析に用いた図の説明と凡例

(1) ワードクラウド

■概要

- ・抽出したデータにおける語句の出現頻度を直感的に把握するための図である。
- ・語句を出現頻度順に、また、出現頻度により書体の大きさを変えて掲載している。

■凡例

- ・赤字：肯定的な意味合いの語句（以下、ポジティブワード）を示している。
- ・青字：否定的な意味合いの語句（以下、ネガティブワード）を示している。
- ・括弧内の数字：その語句が、抽出されたデータのうち何件で用いられているかを示す。

■作図の条件等

- ・Twitter データの分野別分析、コールセンターへの問合せについては、出現頻度が多い順に 150 語を、Twitter データの札幌全体のイメージ分析については 200 語を掲載している。

(2) ワードランキング

■概要

- ・抽出したデータにおける頻出語句を出現頻度順に並べ、抽出したデータにおける出現件数と出現率を掲載している。
- ・Twitter データ分析における〈全体〉〈道内〉〈道外〉の比較と、以下で述べる共起ネットワーク図作成に中心語として用いた語句を示すために用いている。

■凡例

- ・オレンジ、青の下地で示す語句は後で述べる共起ネットワーク図作成において中心語として用いた語句である。
- ・オレンジ：出現件数 10 件以上、出現頻度が概ね上位 15～20 位以内で、当該分野の分析（共起ネットワーク図作成）に有用であると考えた語句。
- ・青：上記条件には当てはまらないが、当該分野の分析に必要と考えた語句。

■作図の条件等

- ・Twitter データの全体イメージ分析については形容詞を、その他については名詞を抽出している。これは、前者についてはイメージを、後者については語られている話題を探るためである。
- ・Twitter データの全体イメージ分析の〈全体〉については出現件数 100 件以上の語句、〈道内〉〈道外〉については 50 件以上の語句を、分野別分析の〈全体〉については、概ね出現件数 10 件以上かつ 150 位までの語句、〈道内〉〈道外〉については概ね出現件数 5 件以上かつ 50 位までの語句を掲載している。

(3) 共起ネットワーク図

■概要

- ・語句間の関連を表すことで、主要な話題、話題間の広がり、つながりについて把握するための図である。
- ・分析は主にこの図によって行った。

■凡例

- ・青色の楕円：中心語。
- ・緑色の楕円：係り受け語。オレンジの線で紐付いている中心語と係り受けの関係で出現する語句。
- ・灰色の矢印：関連度線。同時出現の度合い（共起度）が強い中心語間を結んでいる。片矢印は矢印元の中心語が用いられる際、矢印先の中心語が共起する人が多いことを示す。両矢印は、両方の中心語が同時に共起する人が多いことを示す。線の太さは共起する度合いを示している。
- ・括弧内の数字：中心語については、その語句が抽出されたデータのうち何件で用いられているかを示す。係り受け語については、中心語からの係り受けで何件のデータで用いられたかを示す。
- ・赤字の語句：ポジティブワード。
- ・青字の語句：ネガティブワード。
- ・赤の破線楕円、赤太字：赤の破線楕円内で語られている話題を示す。
- ・青の破線楕円、青太字：上記と同様だが、当該分野には関連しないが注記が必要な話題を示す。
- ・赤太字の中心語：複数語句にはまたがっていないが、当該分野について注目すべき内容が語られた中心語を示す。

■作図の条件等

- ・中心語は、(2) で述べたように、概ね頻出語句を設定した。ただ、「昨日」「今日」、他都市名など、当該分野等の分析にはそぐわない語句は除外している。
- ・中心語に関連した係り受け語については、Twitter データの全体イメージ分析、コールセンターへの問合せの分析については係り受けでの出現数 2 件以上で出現数が多いものから 10 個以内と設定し作図した。Twitter データの分野別分析については、係り受けでの出現数 2 件以上で出現数が多いものから 5 個以内と設定し作図した。

(4) ポジ・ネガグラフ

■概要

- ・<全体><道内><道外>別、性・年代・居住地域別に肯定的（ポジティブ）なツイート、否定的（ネガティブ）なツイートの比率を示したグラフ。
- ・当該分野等に関する属性間の受け止め方の差異を見るために用いた。

■凡例

- ・N：それぞれの属性について、ポジティブ・ネガティブが判別されたツイートの件数。

■作図の条件等

- ・ポジティブワードが使われているツイートをポジティブ、ネガティブワードが使われているツイートをネガティブとしてカウントしている。

(5) 出現率グラフ

■概要

- ・ <全体><道内><道外>別、性・年代・居住地域別に、その分野に関するツイート、コールセンターへの問合せがどれだけの比率で出されているかを示したグラフ。
- ・ 当該分野等に関する属性間の関心の差異を見るために用いた。

■凡例

- ・ N：それぞれの属性の総件数。

5. 分析についての注記

(1) 共起ネットワーク図の分析について

- ・ 上述したように、Twitter における話題、コールセンターへの問合せ内容の分析は主にこの図によって行った。
- ・ Twitter の分析については、まず<道内>と<道外>別に分析した内容を記述し、<全体>については、<道内><道外>で語られていない話題が見られた等の場合に補足的に記述している。
- ・ 出現頻度が高い中心語、共起関係、係り受け関係が見られても、それらの語句が、分析する分野等とは関係の薄い話題の中で使われている場合、札幌と関係のない話題の中で使われている場合については記述していない。
- ・ 取り上げたコールセンターへの問合せについては、代表的な問合せ内容を、性・年代・居住地とともに掲載している。

(2) ポジ・ネガグラフ、出現率グラフの分析について

- ・ これらは、当該分野等に関する属性間の差異を見るために用いた。
- ・ 差異については、母比率の差の検定（Z 検定）を行い、有意水準 5%にて有意差が見られたものについてのみ記述している（一見差異があるように見えても、有意差がなかったものについては記述していない）。