

-2020年度学生研究奨励賞-

自動車顧客購買データを用いた 新車販売店舗における従業員の接客行動分析

東海大学 情報通信学部 経営システム工学科

朝日研究室 山岸大祐

目次

- 1.研究背景
 - 2.研究対象
 - 3.先行研究
 - 4.研究目的・手順
 - 5.データ概要
 - 6.データ項目
 - 7.データ処理
 - 8.分析に使用するソフトウェア
 - 9.分析手順
 - 10.基礎集計
 - 11.相関分析
 - 12.相関分析結果
 - 13.非階層型クラスター分析
 - 14.クラスター分析結果
 - 15.接客行動内容の分類
 - 16.重回帰分析
 - 17.各クラスター毎の保有顧客年代
 - 18.各クラスター毎の販売車種
- まとめ&今後の課題
- 参考文献
- Appendix

研究背景-自動車業界の現状-

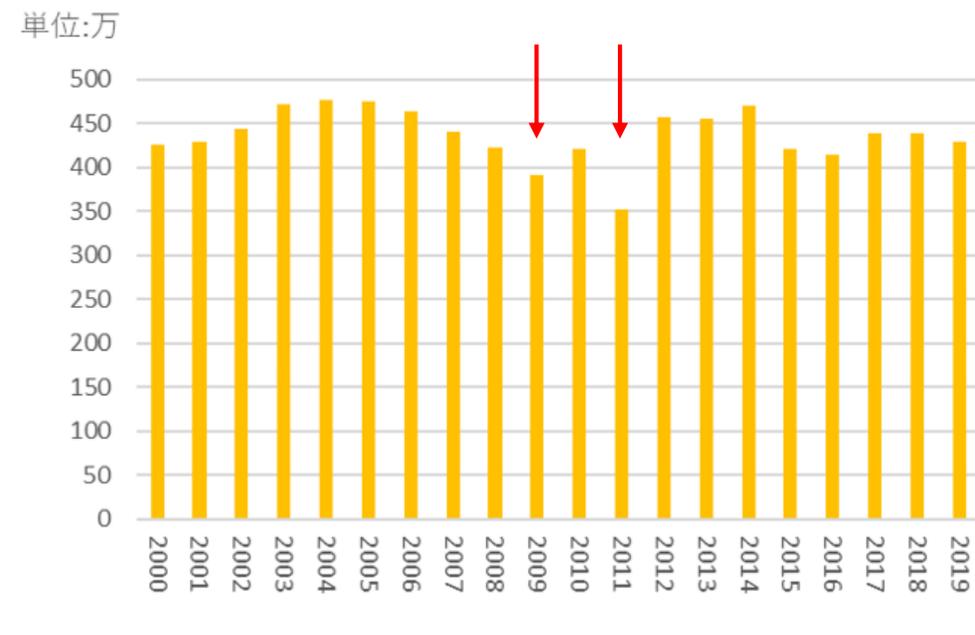


図1,新車乗用車販売台数の年度別推移

▷新車販売台数はここ5年間では横ばい傾向

売上が大きく下降した2009年はリーマンショック
2011年は東日本大震災が起きた年である

今年2020年においては世界的に新型コロナウイルス
感染が拡大,日本でも緊急事態宣言が出された

▷売上が下降した2つの時期と同様
今年度においても消費者購買行動が冷え込む可能性
が高いと考えられる

▷自動車産業においても影響を受けることを鑑み
販売台数維持の為の販売戦略を計画する必要がある
(年)

研究目的

新車販売台数の維持・向上

研究背景-地域毎の乗用車に対する利便性-

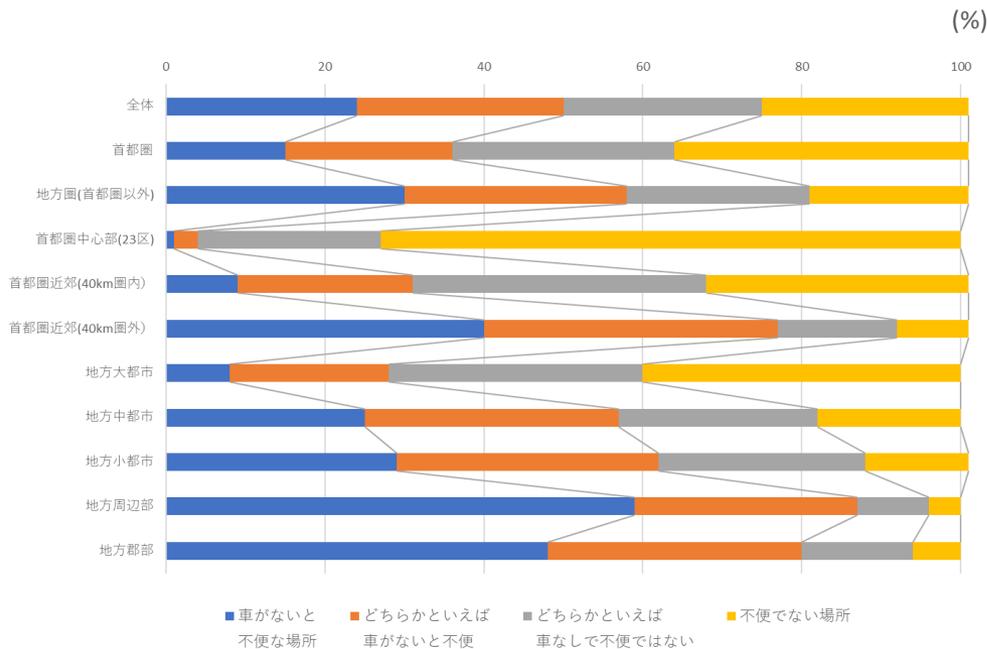


図2,地域別乗用車の利便性調査結果

▷図2は地域別毎の世帯を対象に乗用車の利便性について調査した結果である

▷この結果より首都圏中心部や地方大都市といった場所に住む世帯は,車を交通手段としての利便性を感じる世帯が少ない一方で,規模が小さい都市では乗用車がないと不便と感じる世帯が多いと推測できる

▷従って首都圏近郊部(40km圏外)地域における業務内容の改善によって新車販売台数の向上が見込める

研究対象

→従業員と顧客、どちらを研究対象にするか？

研究対象:顧客



図3,顧客を対象にした場合

- ▷顧客が購入する商品や自動車に求めるニーズを題材にした研究となる
- ▷顧客のニーズは時間と共に絶えず変化する為研究を長期で活用するのは難しい,さらに企業単位での企業戦略の為 慎重かつ大規模な計画が必要

研究対象:従業員(営業職)



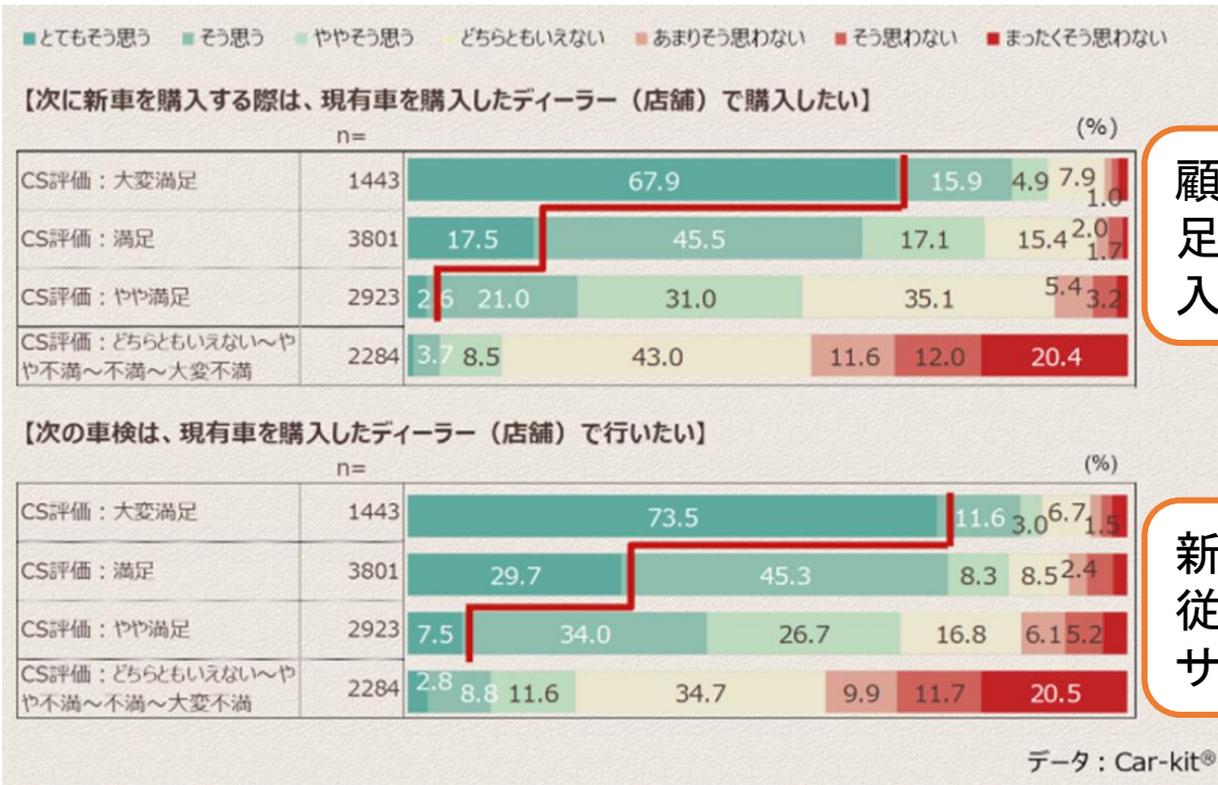
図4,従業員を対象にした場合

- ▷従業員の労働内容や顧客に対する接客サービスの研究となる
- ▷従業員個人がノウハウとして研究を長期で活用できる為,自身で研究内容を基に接客サービスの改善と発展への取り組みができる

研究対象

営業職の
従業員

研究対象-従業員を対象にするメリット-



顧客が従業員の接客に対する満足度が高い程同じ店で新車を購入したいと強く思う傾向にある

新車販売台数を向上させる為に従業員の接客業務改善活動やサービスの向上が必要である

図5.顧客満足度で分類した今後のサービス意向

(インテージ調べ)

先行研究

▷松山(2019)の卒業研究

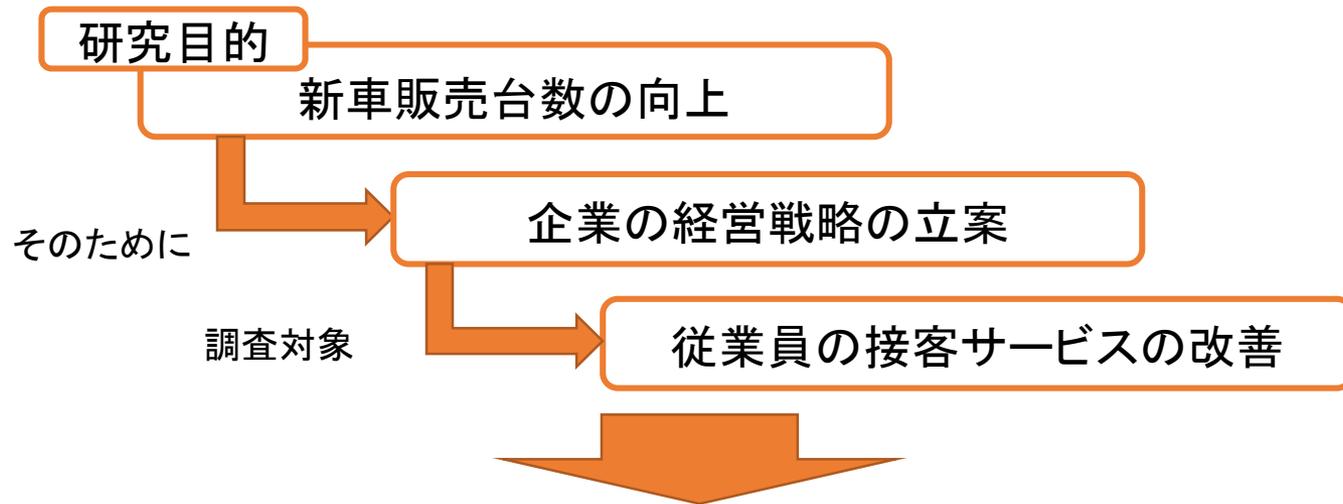
→新車販売台数向上を目的とした研究

▷営業職の従業員の社内評価に疑問を抱き分析によって従業員を再評価し優劣の従業員それぞれの保有顧客を分析しどのような顧客を増やせば売上が向上するか調査した研究

↳ 従業員の評価基準が新車販売台数の一つであった

本研究では従業員を評価するにあたって顧客との接客の内容も併せて分析し従業員を評価その分析結果を用いて新車販売台数向上の為の提案を行う

研究目的・手順



▷新車販売業務を行う従業員の顧客との接客頻度に注目

従業員を**優秀**な従業員と**伸び悩む**従業員に分類する

評価基準は販売台数と顧客との対面での接客頻度の2要素

優秀な従業員と**伸び悩む**従業員それぞれの保有顧客と顧客に対する接客行動の特徴を発見し**伸び悩む**従業員に対する改善案を検討する

データ概要

▷データ概要:四輪自動車販売会社A

データ提供元:ファイナンス事業会社B

集計期間:2017/04/01～2019/03/31

対象地域:特定地域内(首都圏近郊部40km圏外)における新車販売店舗

▷データ内容

顧客購買記録データ...新車購入者の性別,年齢等

従業員接客記録データ...顧客との接触時間,具体的な行動内容等

顧客乗用車メンテナンス記録 等

▷データ数

全従業員数:211人

全保有顧客データ数:11,915人

全顧客接触履歴データ数:2,574,070個

データ項目 -使用したデータ項目-

表1,使用したデータ項目

説明	形式
顧客の固有ID	6から9桁の整数
営業社員の固有ID	5から6桁の整数
店舗の固有ID	1から2桁の整数
従業員の所属店舗名	文字列
新車購入契約の個別ID	7桁の整数
顧客が新車購入契約を行った日付	日付
従業員が顧客に接客行動を行った日付	日付
顧客の性別(男:1,女:2,その他不明:3)	1桁の整数
顧客の年齢	2桁の整数
顧客が購入した乗用車の名称	文字列
従業員が顧客と接触した回数	1から4桁の整数
従業員毎の新車売上台数	1から2桁の整数
顧客との接触手段	文字列
接客相手	文字列
社員の氏名	文字列

データ処理

▷集計期間:2017/04/01～2019/03/31

↳ ▷期間を2018/04/01～2019/03/31の1年間に設定

▷接客相手:法人を分析対象外に設定

↳ 従業員の得意先である傾向がある為、接客行動による評価が難しいため

▷接客記録内の従業員の接客相手を本人のみを対象とする

↳ 新車購買契約をする際に直接顧客と対面しないで完了するとは考えにくいため

▷接客行動記録内の「訪問,電話,メール」を分析対象外に設定

↳ 留守電や顧客が家にいない状況での訪問のデータが含まれている為

▷データ処理後

従業員数: 211人→157人 顧客データ数:11,915人→2,971人 全顧客接触履歴データ数:2,574,070個→257,205個

分析に使用するソフトウェア

▷分析には株式会社NTTデータ数理システム様が提供されているVMSver.8.7 (Visual Mining Studio)を使用させて頂いた

VMSとはデータマイニングソフトウェアであり多様な形式のデータを取り込むことができかつ大量のデータを扱う事ができる

VMSを分析で用いる利点として分析の流れをフローで追うことができる点が挙げられる

データ分析を再度実行したい場合においてもフロー元を辿り、再実行することができるため、1からやり直す手間を削減し、効率的にデータ分析をすることができる

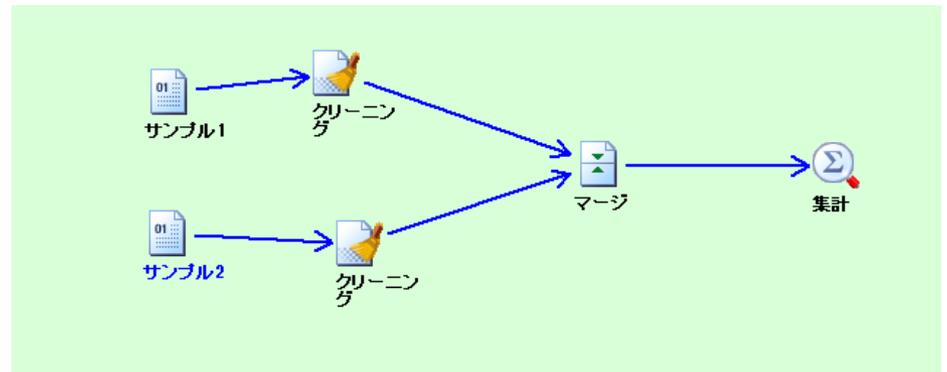


図6.分析時の操作画面例

分析手順

- ✓ 1.基礎集計によって顧客の特性を把握→提案で活用
- ✓ 2.相関分析で接客回数と販売台数に相関はあるかを確認
- ✓ 3.クラスター分析で従業員の優劣を分類
- ✓ 4.優劣の従業員毎に接客行動の差を接客行動を分析することにより発見する

基礎集計 -売上上位の車種-

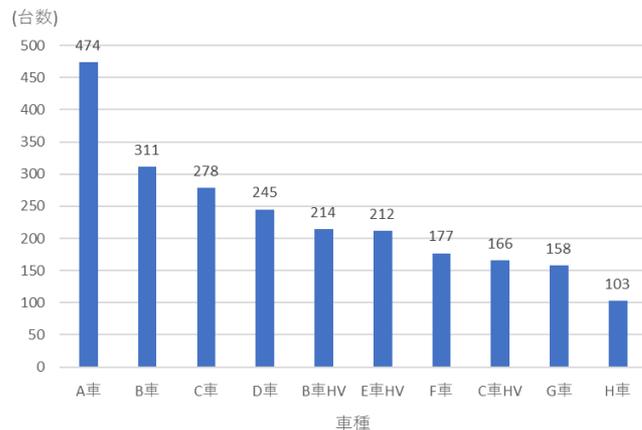


図7,一年間における新車販売台数上位十種

表2,新車販売台数上位十種の性能表

車名(仮称)	車名.数	乗車人数	燃費	価格(参考)
A車	474	5	24.6km/L	¥1,556,500
B車	311	7	13.2km/L	¥2,844,600
C車	278	5	25km/L	¥1,509,200
D車	245	5	27.2km/L	¥1,921,700
B車HV	214	7	19km/L	¥3,347,300
E車HV	212	5	25.8km/L	¥3,045,000
F車	177	5	30.8km/L	¥2,655,000
C車HV	166	5	34.4km/L	¥1,853,500
G車	158	7	10.8km/L	¥4,539,000
H車	103	5	17km/L	¥2,098,000

多くの顧客に購入される新車は大きく二種類に分類できる

- ▷低燃費・低価格
- ▷乗車人数5～7人のファミリーカー

このことから自社で抱える顧客に車に安価で高性能を求める顧客と家族世帯持ちの顧客がいることが推測できる

※HV=ハイブリッド車のこと

基礎集計 -購買者の年齢-

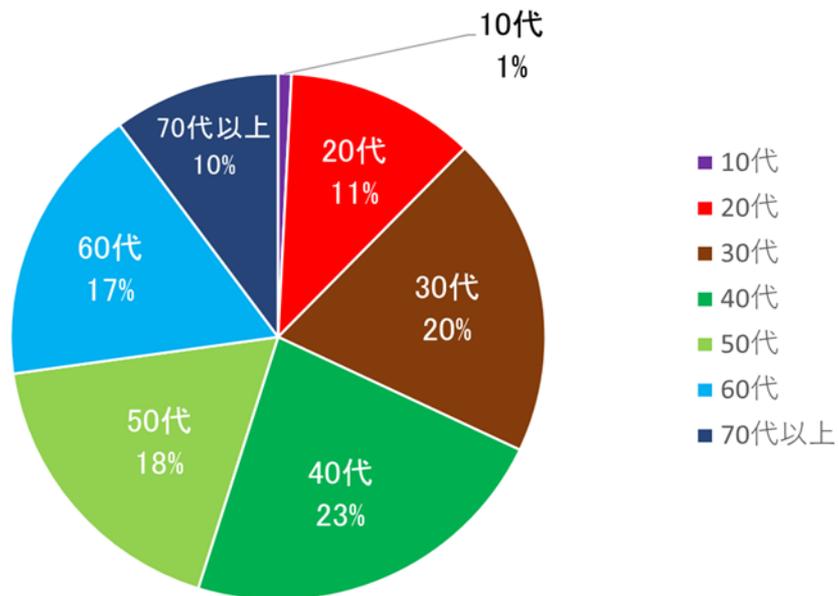


図8.新車購入者の年代別割合

▷新車購入者は30代から40代の顧客が全体の4割弱を担っている

▷70代以降新車購入者が減少するのは高齢者となった利用者が免許返納を行う等して車に乗る機会が減少するためと推測する

基礎集計 -国内配偶者率-

▷男女共通して30代を起点に配偶者がいる割合が未婚率よりも上回っている
それ以降も50代半ばまで割合が向上傾向にある

▷自動車販売店においても、来店する顧客は30代～40代が中心であることやファミリーカーが売り上げ上位十種に属していることから、
新車の購入を計画する顧客は配偶者のいる家族世帯が多い可能性が高い

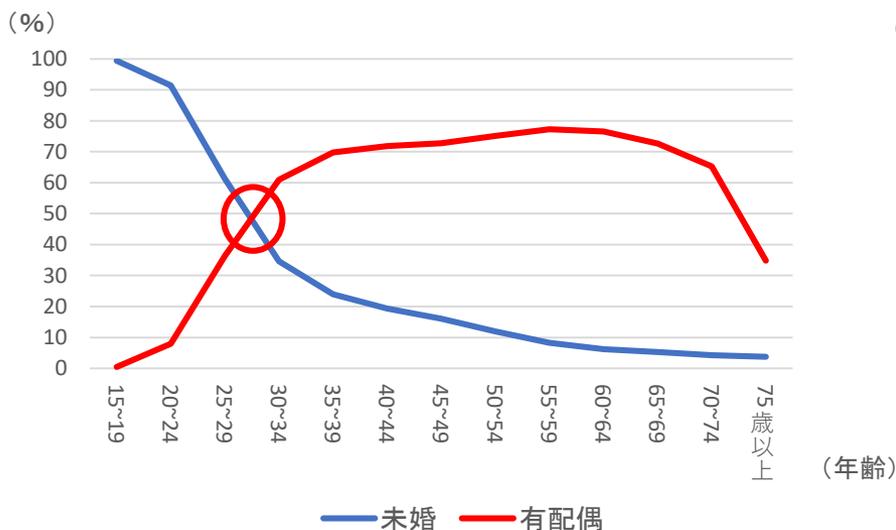


図9,男性の年齢別配偶者率推移

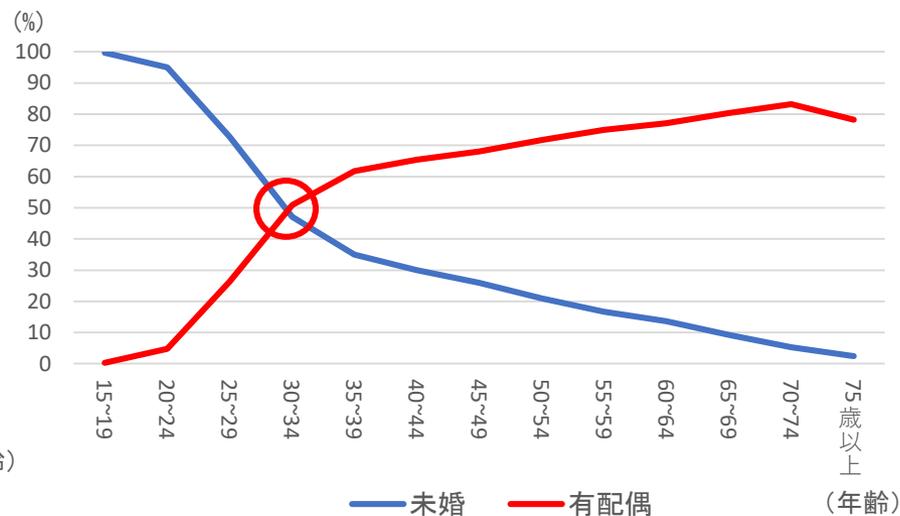


図10,女性の年齢別配偶者率推移

基礎集計 -国内配偶者率-

▷この結果から新車を購入する顧客の割合は、一週間の中で祝日にあたる土曜日と日曜日が高い

▷祝日の割合が高い要因として新車購入を家族全員で来店し意思決定を行えることが考えられる夫婦が共働きまたは片働きであった場合でも祝日なら一般的に勤務先が休日であるため家族で店舗に来店することができる

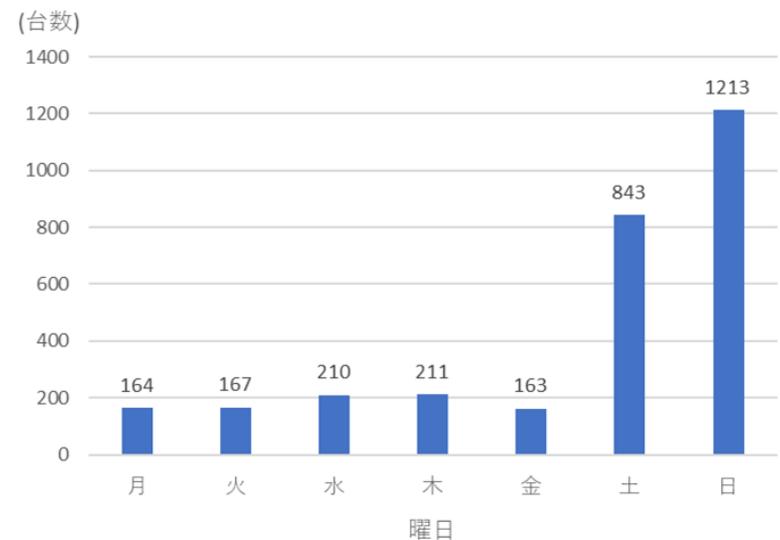


図11.男性の年齢別配偶者率推移

相関分析

- ▷データ内における二つの要素に相関関係があるのかを調べる
相関とは2つの要素を散布図で示した場合、右上がりまたは左下がりで表すことができる関係にあることである
- ▷相関分析はこの関係性を「相関係数」と呼称される数値で表現する分析手法である
- ▷本研究においては顧客購買データ内で集計した新車販売台数と従業員が顧客と接客した回数を集計した接触回数の2つを要素として用いる
- ▷相関分析を行うにはVMS内に存在する相関アイコンを使用する



図12.相関アイコン

相関分析結果

- ▷分析の結果,従業員と顧客との接客回数と新車販売台数には強い正の相関関係が見られた
- ▷接客と顧客の購買行動には関係性はあると考えられる
この結果をふまえ次の分析手法を用いて接客と売上の関係を従業員評価と共に検証する
- ▷相関分析をする都合上,接客回数を従業員の顧客に対して接客した合計で検討している
- ▷相関分析の結果をふまえ上で従業員が顧客一人に行った従業員の接客回数と新車販売台数を基準に次の非階層型クラスター分析を用いて評価付けを行う

	販売台数(平均)
接客回数(平均)	0.87

図13,相関分析結果

非階層型クラスター分析

▷従業員を優秀な従業員と売上の伸び悩む従業員に分類するにあたって非階層型クラスター分析を用いる この分析によりデータが似ている物同士で分類を行う

クラスター分析とは多量なデータ集団内にある似ているデータ同士を集めグループに分類するデータ分析手法である

▷K-means法とはクラスター分析手法におけるデータ間の距離計測方法の一つであり、クラスターの平均を計測、その結果からクラスターの分類数を決定する
尚欠点として初期値依存が高い為複数回分析を試行した

▷クラスター分析を行うにあたってVMS内のK-means法アイコンを使用する

計測設定

初期クラスターの設定方法:k-means法

距離の計測方法:Euclid法

表示方法:対散布図

各クラスターの数値を比較して名前付け



図14,K-Means法アイコン

クラスター分析 -分類数の決定-

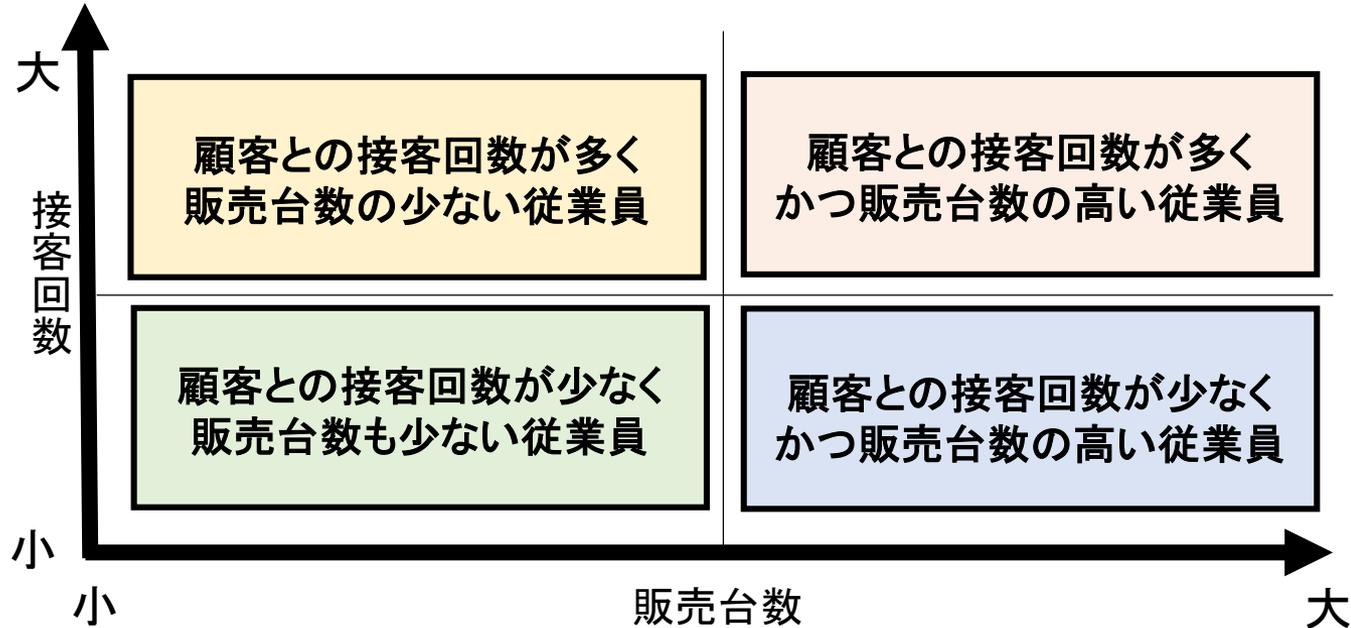


図15.クラスター分析分類モデル図

- ▷ クラスター分析は従業員の新車販売台数と接客回数の2要素を用いる
- ▷ 分類数を4に設定することによって販売台数の大小、接客回数の大小の組み合わせ全通り(4通り)を検討することができる

クラスター分析結果

▷ クラスターID1 接客回数:小 販売台数:大
→ 労働量を抑えて販売することができる従業員である
↳ **優秀**な従業員に設定

▷ クラスターID2 接客回数:大 販売台数:小
→ 労働量に対して販売台数の少ない従業員である
↳ **伸び悩む**従業員に設定

表3, クラスター結果

ID	保有顧客数	接客回数(平均)	販売台数(平均)	クラスターに属する従業員数
1	647	8.64	46.72	16
2	737	7.2	13.18	94
3	220	16.41	63.51	10
4	1312	6.69	28.92	48

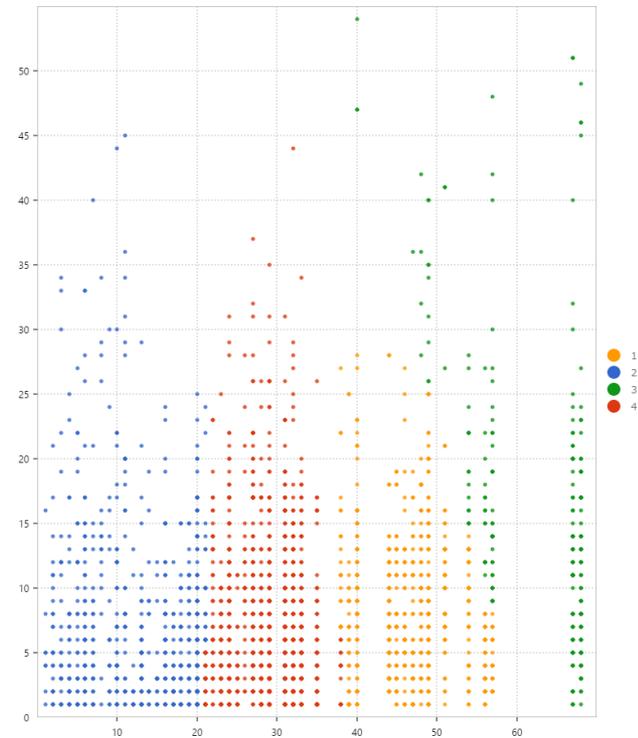


図16, クラスター分析結果散布図

接客行動内容の分類

- ▷従業員の接客行動内容について検討する際、顧客に対しての接客活動の種類は膨大である。従ってそのままの接客活動データではクラスター毎の従業員の特徴を視覚化することは難しい。
- ▷そのため接客行動を新車の購買に関連性がある項目と車検や修理といった新車購入には関連性のない行動で区分する。

本研究では接客行動を以下の四つに分類し、各従業員毎に四つの行動を集計した。

▷接客活動区分

- ✓ 1.実際に新車を販売したあるいは入庫したかがわかるデータ(自店社内入庫確認済等)
- ✓ 2.新車販売に関わる接客活動途中の記録(入庫商談中,カタログ渡し等)
- ✓ 3.その顧客に対する新車販売接客活動を断念した記録(活動断念・敗戦等)
- ✓ 4.それ以外の新車販売や入庫には関係性のない記録

重回帰分析

- ▷接客行動を四つに区分したデータを使用し重回帰分析を行う
回帰分析とは説明変数と呼ばれる原因を元に目的変数という結果を予測するデータ分析手法であり
その中でも説明変数が複数ある場合が重回帰分析である
- ▷回帰分析結果は切片と説明変数に係っている傾き回帰係数で表される回帰式という数式で出力される
- ▷本研究では従業員の販売台数を目的変数に設定,説明変数をp25で述べた従業員の接客活動区分四つに設定する
- ▷重回帰分析を行うにあたってExcelのアドインである「分析ツール-VBA」を使用した

回帰式: =切片+回帰係数×説明変数

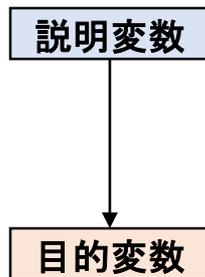


図19,回帰分析

回帰式: =切片+偏回帰係数A×説明変数A
+偏回帰係数B×説明変数B+偏回帰係数C×説明変数C

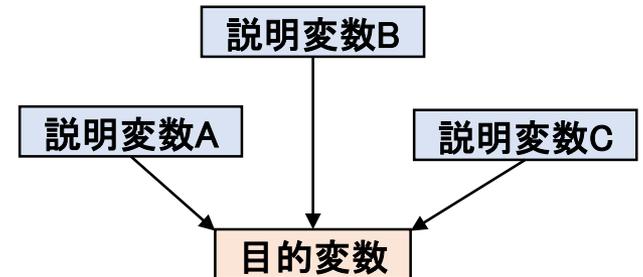


図20,重回帰分析

重回帰分析-結果-

▷クラスター1の従業員データを用いた重回帰分析結果

回帰統計	
重相関 R	0.626703707
重決定 R ²	0.392757536
補正 R ²	0.171942094
標準誤差	5.354900113
観測数	16

分散分析表

	自由度	変動	分散	観測された分散比	有意 F
回帰	4	204.0129926	51.00324814	1.778668798	0.203330109
残差	11	315.4245074	28.67495522		
合計	15	519.4375			

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	42.54191032	3.444464393	12.35080566	8.64605E-08	34.96069531	50.12312534	34.96069531	50.12312534
活動区分1	0.004594413	0.010785352	0.425986344	0.67833365	-0.019143988	0.028332813	-0.019143988	0.028332813
活動区分2	0.013737991	0.019695282	0.69752697	0.499950883	-0.029611033	0.057087015	-0.029611033	0.057087015
活動区分3	-0.011036324	0.017971435	-0.614103633	0.551636005	-0.050591185	0.028518538	-0.050591185	0.028518538
活動区分4	0.002553276	0.015096463	0.16913073	0.868763868	-0.030673816	0.035780368	-0.030673816	0.035780368

有意Fが0.05未満でないため
帰無仮説を棄却できない

接客活動区分を用いた重回帰分析では
従業員の特徴を確かめることはできなかった

各クラスター毎の保有顧客年代

- ▷ 従業員の特徴を発見するために各クラスターにおける従業員の保有顧客に注目する
- ▷ 図17,18はクラスター1,2毎における顧客の年代別割合である
- ▷ p14-15で示したように主な顧客は30~40代の家族世帯である
優秀な従業員は伸び悩む従業員より主要顧客である30~40代の購買顧客の割合が高い

▷ クラスター1

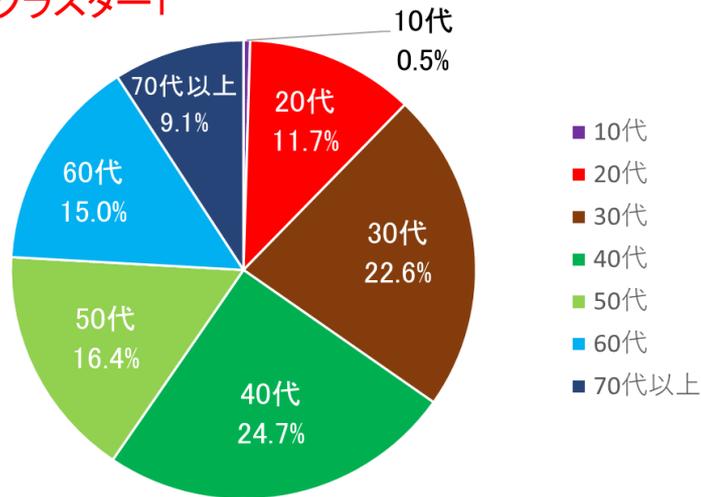


図17,優秀な従業員の購買顧客年齢分布

▷ クラスター2

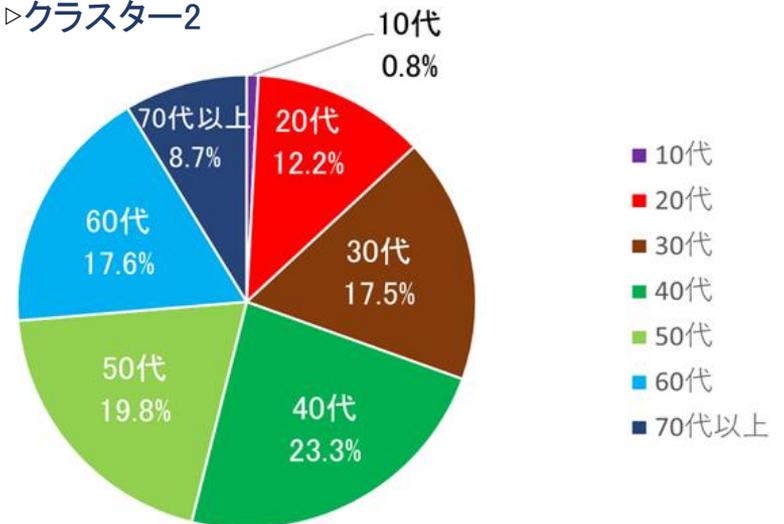


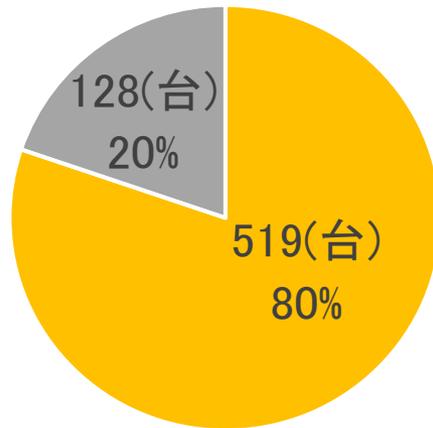
図18,伸び悩む従業員の購買顧客年齢分布

各クラスター毎の販売車種の特徴

- ▷ 図19,20はクラスター1,2毎の販売車種における売上上位十種の割合である
 - ▷ 優秀な従業員の販売した車種は伸び悩む従業員より売上上位十種の割合が高い
- ↳ P23で述べた保有顧客年代と併せて考えると優秀な従業員は

来店する主要な顧客に対して売れ筋商品を販売し店舗売上を伸ばしている

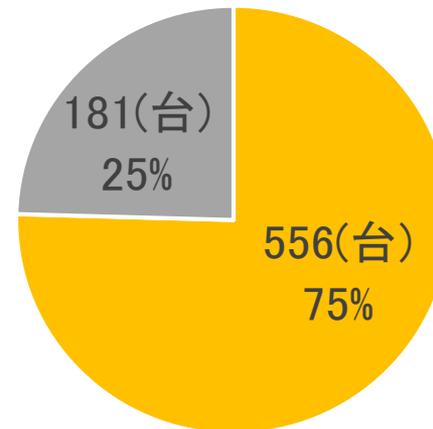
▷ クラスター1



■ 売上上位10車 ■ その他車種

図19,優秀な従業員の販売車種の売上上位十種の割合

▷ クラスター2



■ 売上上位10車 ■ その他車種

図20,伸び悩む従業員の販売車種の売上上位十種の割合

※売上上位十種はp14に記載

まとめ

▷相関分析

従業員の接客回数と新車販売台数の関係性について調査
この結果より相関係数は0.87となり強い相関関係がみられた

▷非階層型クラスター分析

従業員を優秀な従業員と売上の伸び悩む従業員に分類
接客回数と新車販売台数の大小四つから分類数を4に決定
クラスターの数値特徴から優秀な従業員と伸び悩む従業員を決定

▷クラスター毎の従業員の接客活動集計

優秀な従業員は伸び悩む従業員よりも対象地域店舗に来店する顧客の主要な年代の顧客に対して売れ筋商品を販売している特徴がみられた

▷今後の課題

活動結果区分を用いた重回帰分析ではあまり良い結果が出なかった
活動結果区分を用いた分析から優秀な従業員の接客行動について検討していきたい

参考文献

[1]カーディーラーに求められるサービスとは？顧客満足度(CS)調査から見えた重要ポイント

{<https://www.intage.co.jp/gallery/car-cssurvey/>}

(最終閲覧日:2020/08/08)

[2]NHK:新型コロナウイルス特設サイト

{<https://www3.nhk.or.jp/news/special/coronavirus/data-all/>}

(最終閲覧日:2020/08/08)

[3]一般社団法人 日本自動車工業会:アクティブマトリックスデータベースシステム

{<http://jamaserv.jama.or.jp/newdb/index.html>}

(最終閲覧日:2020/08/08)

[4]一般社団法人 日本自動車工業会:2019年度乗用車市場動向調査

{http://www.jama.or.jp/lib/invest_analysis/pdf/2019PassengerCars.pdf}

(最終閲覧日:2020/08/09)

[5]松山芳生:自動車業界における営業優劣分類方法および顧客属性分析,

東海大学 情報通信学部経営システム工学科卒業論文(2019)

[6]平成27年度国勢調査,配偶関係,年齢,(歳階級),男女別 15歳以上人口の割合

{<https://www.stat.go.jp/data/kokusei/2015/kekka/kihon1/pdf/gaiyou1.pdf>}

[7]松井剛,西川英彦:1からの消費者行動,碩学舎,第11章家族,pp160(2020)

[8]恩藏直人,富田健司:1からのマーケティング分析,碩学舎,第7章相関分析,pp112-121(2011)

[9]小野田崇,坂井美帆,山田誠二:K-means法の様々な初期値設定によるクラスタリング結果の実験的比較,(財)電力中央研究所,東京工業大学大学院,国立情報学研究所(2011)

{https://www.jstage.jst.go.jp/article/pjsai/JSAI2011/0/JSAI2011_1J1OS91/_pdf}

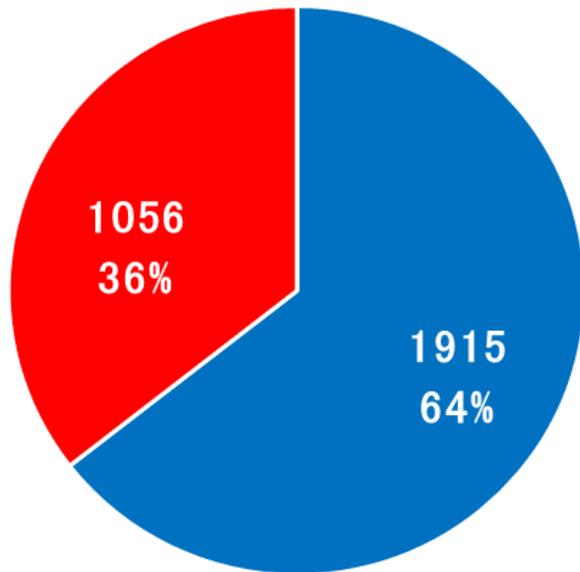
(最終閲覧日:2020/10/30)

参考文献

[10] 恩藏直人, 富田健司: 1からのマーケティング分析, 碩学舎, 第7章 相関分析, pp177-181 (2011)
(最終閲覧日: 2020/12/09)

Appendix

基礎集計



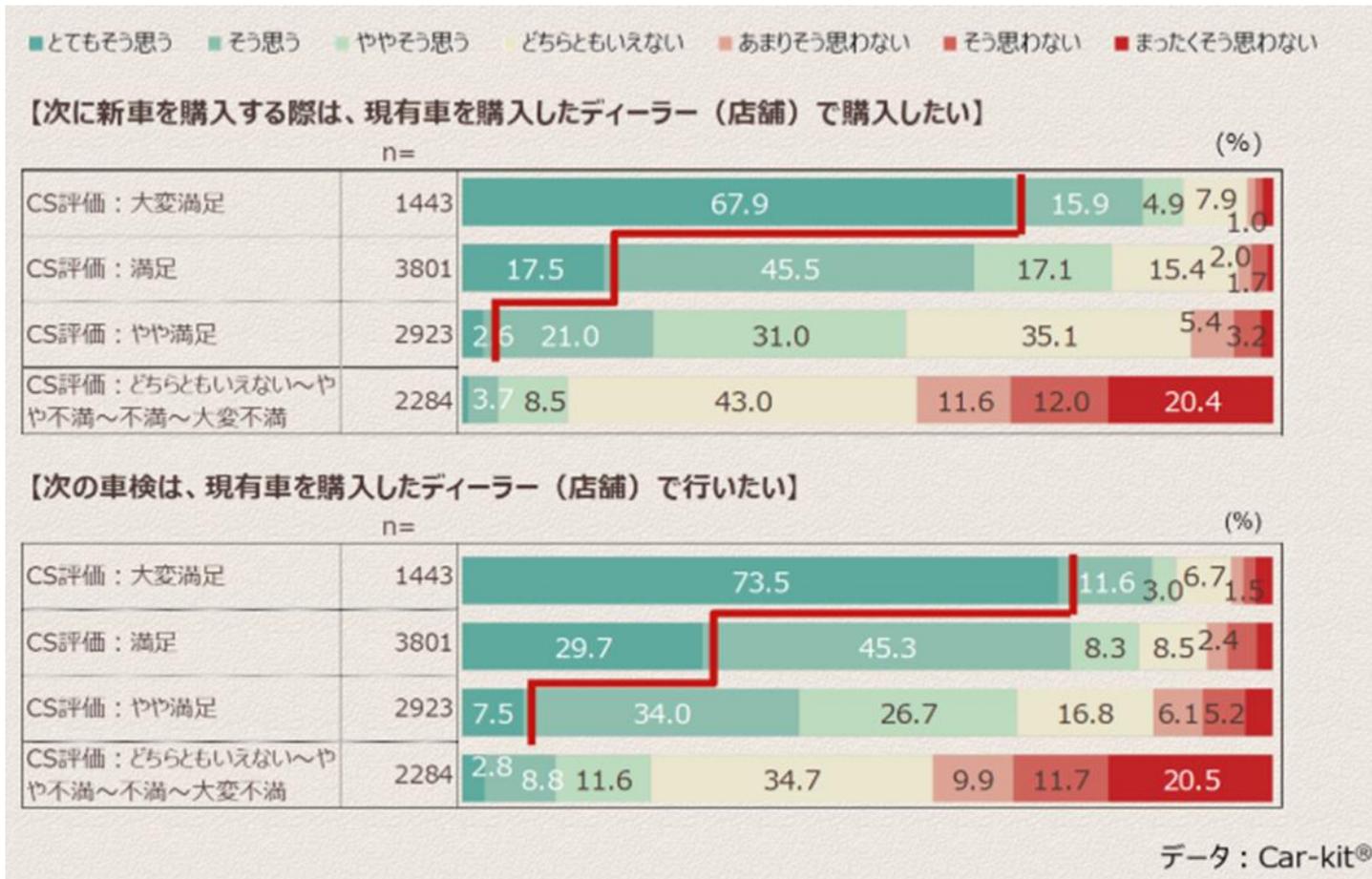
■ 男性
■ 女性

▷新車購入者の割合は男性が6割女性が3割という結果となり、男性の購入者が過半数を超えている

しかし世帯主が購入する場合、夫の名義で新車購入し夫婦ともに自動車を利用する可能性も棄却できないため参考程度に留めておく

図10,新車購入者の性別割合

顧客満足度と購入意思の関係



(インテージ調べ)

相関分析

▷相関分析とはデータ内における二つの要素に相関関係があるのかを調べる分析である
ここで述べた相関とは2つの要素を散布図で示した場合に右上がりまたは左下がりの直線で表すことができる関係性にあることである

▷相関分析はこの関係性を共分散の持つもともとの尺度の大小による影響を取り除き
関連性の強さだけを表現した「相関係数」で示す分析手法である

共分散とはデータ同士の関連性が強いほど絶対値が大きくなり関連性が弱いほど絶対値が0に近づく性質を持つ数値である

共分散を求める計算式 S_{xy} は(3.1)で示される

$$S_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

相関分析の数式

クラスタ分析 - K-means法 -

▷K-means法 アルゴリズム手順

- 1.任意のk個のクラスタ中心 c_i を一様ランダムに選択する
- 2.全てのデータを各データ点 $x_j, j \in \{1, \dots, n\}$ から最も近いクラスタ i に割り当てる
- 3.各クラスタごとに,以下数式に従ってクラスタ中心を求める

$$c_i = \frac{1}{|C_i|} \sum_{x_j \in C_i} x_j$$

C_i は各クラスタ i に含まれるデータの集合であり $|C_i|$ はクラスタ C_i に含まれるデータ数である

- 4.クラスタに変化がなくなるまで手順2.3を繰り返す

クラスタ分析 - K-means++法 -

▷K-means++法 アルゴリズム手順

a, 一つ目のクラスタ中心 c_1 をデータ X からランダムに選択する

b, 全てのデータに対して $D(x_j), j \in 1, \dots, n$ を求める

ここでの $D(x_j)$ とはデータ点 x_j と既に設定されたクラスタ中心との最短距離を表す

c, 次式を満たす実数値 L をランダムに求める

$$0 \geq L \geq \sum_{j=1}^n D(x_j)^2$$

d, 次式を満たす実数値 x_j を次のクラスタ中心 c_l に選択する

$$\sum_{j=1}^{l-1} D(x_j)^2 \geq L \geq \sum_{j=1}^l D(x_j)^2$$

e, クラスタ中心を k 個選ぶまで手順bとdを繰り返す

k 個選択後K-means法のアルゴリズムの手順2.3.4と同様の処理を行う

重回帰分析に使用したデータ形式

ID	売上台数	活動区分1	活動区分2	活動区分3	活動区分4
1	40	765	538	573	1308
2	51	1110	435	235	956
3	49	777	931	1445	2431
4	56	718	314	156	889
5	49	468	171	339	749
6	44	240	241	334	673
7	38	514	178	387	891
8	40	299	141	536	957
9	54	631	368	368	1139
10	57	1161	539	579	1530
11	48	386	381	720	1054
12	45	79	142	176	390
13	46	283	317	469	850
14	44	218	135	342	654
15	47	179	241	519	680
16	39	99	108	151	371

接客回数を集計方法

▷ 相関分析に使用した接客回数を集計方法

