

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



پردیس البرز

## عنوان پایان نامه یا رساله

تجزیه و تحلیل کاربرد Big Data در فرآیندهای بازاریابی

نام دانشجو

شبنم جلالی

استاد راهنما: جناب آقای دکتر اصلانی

استاد مشاور: جناب آقای دکتر نظری

پایان نامه برای دریافت درجه کارشناسی ارشد

دی ماه سال ۹۶



پدیس البرز

گروه آموزشی مدیریت

گواهی دفاع از پایان نامه کارشناسی ارشد

هیأت داوران، پایان نامه کارشناسی ارشد: خانم شبنم جلالی به شماره دانشجویی 150294239 در رشته

مدیریت بازرگانی گرایش بازاریابی در تاریخ 96/11/01

با عنوان: "تجزیه و تحلیل کاربرد Big Data در فرآیندهای بازاریابی"

نمره نهایی:

به عدد: ۱۷ به حروف: هفده

و درجه: خوب ارزیابی نمودند.

ردیف	مشخصات هیأت داوران	نام و نام خانوادگی	مرتبه دانشگاهی	دانشگاه یا مؤسسه	امضاء
۱	استاد راهنما	دکتر علیرضا اصلانی	استادیار	دانشگاه تهران	
۲	استاد مشاور	دکتر محسن نظری	دانشیار	دانشگاه تهران	
۳	استاد داور داخلی	دکتر سعید روحانی	استادیار	دانشگاه تهران	
۴	نماینده تحصیلات تکمیلی	دکتر محمد رحیم اسفیدانی	استادیار	دانشگاه تهران	

تذکر: این برگه پس از تکمیل توسط هیأت داوران در نخستین صفحه پایان نامه درج گردد.

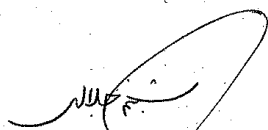
## اداره کل تحصیلات تکمیلی

### تعهدنامه اصالت اثر

اینجانب سید مجتهد می شوم که مطالب مندرج در این پایان نامه/رساله حاصل کار پژوهشی اینجانب است و به دستاوردهای پژوهشی دیگران که در این پژوهش از آنها استفاده شده است، مطابق مقررات ارجاع و در فهرست منابع و مآخذ ذکر گردیده است.

این پایان نامه/رساله قبلاً برای احراز هیچ مدرک هم سطح یا بالاتر ارائه نشده است. در صورت اثبات تخلف (در هر زمان) مدرک تحصیلی صادر شده توسط دانشگاه از اعتبار ساقط خواهد شد.

کلیه حقوق مادی و معنوی این اثر متعلق به پردیس/دانشکده/مرکز دانشگاه تهران می باشد.



نام و نام خانوادگی دانشجو

امضاء

۹۷/۲/۲



بسم الله الرحمن الرحيم

## سوگند نامه دانش آموختگان کارشناسی ارشد دانشگاه تهران

سپاس یزدان پاک را که مرا مشمول الطاف خویش نمود که با طی مراحل تحصیل به اخذ درجه کارشناسی ارشد نائل شوم به شکرانه این نعمت بزرگ الهی که با امکانات این مرز و بوم فراهم و نزد اینجانب به امانت گذاشته شده است، در پیشگاه ملت ایران به کتاب آسمانی خود، قرآن کریم، سوگند یاد می کنم که:

✓ در سراسر زندگی حرفه ای، به نحو احسن در راه اعتلای کشور ایران و جامعه بشری قدم برداشته و در این راه از هیچ کوششی دریغ نکنم.

✓ در تمام فعالیت های تخصصی، رضای خدا را همراه با صداقت علمی و اجتماعی مد نظر داشته و از موقعیت های به دست آمده در جهت رفع مشکلات مردم استفاده کنم و در همه امور، منافع کشور را بر منافع فردی مقدم بدارم.

✓ همواره علم و دانش خود را به روز نگاه داشته و در ادای وظایف و تعهدات حرفه ای در حد توان سعی و تلاش خود را به کار گیرم.

✓ و اینک از خداوند متعال توفیق بندگی و پای بندی به مفاد این سوگندنامه را خواستارم و از او می خواهم که مرا در ادامه و پیمودن مسیر و فتح قله های رفیع علم و دانش و ایفای رسالت علمی و انسانی خویش موفق بدارد.

و ایمان دارم که:

**"إِنَّ اللَّهَ يَعْلَمُ غَيْبَ السَّمَاوَاتِ وَالْأَرْضِ وَاللَّهُ بَصِيرٌ بِمَا تَعْمَلُونَ" (سوره حجرات، آیه ۱۸)**

نام و نام خانوادگی دانشجو امضاء

چکیده:

در دنیای پیچیده امروزی که مدام در حال تغییر و تحول حوزه های گوناگون فناوری و تکنولوژی می باشد ایجاد ظرفیت ها و توانایی های لازم جهت حفظ فضای کسب و کار و رقابت با سایر رقبا به امری کاملاً ضروری برای بقا تبدیل گردیده است. با در نظر گرفتن شرایط حاکم بر بازارهای امروزی کاملاً می توان متوجه شد که دریافت و شناخت نیازهای دائماً در حال تغییر مشتریان و مردم مهم ترین فعالیت و اقدامات شرکت های فعال بازار را شکل می دهد چرا که هر کدام از این شرکت ها تنها در صورتی می توانند به حیات خود ادامه دهند که خود را با این تقاضا ها هماهنگ ساخته و گوی رقابت را از سایر رقبا برابند. با استفاده از تحلیل حجم های بیشتری از داده ها، می توان تحلیل های بهتر و پیشرفته تری را برای مقاصد مختلف، از جمله مقاصد تجاری، پزشکی و امنیتی، انجام داد و نتایج مناسب تری را دریافت کرد. هدف این پژوهش تحلیل کاربرد Big Data بر فرآیند های بازاریابی و رقابت می باشد. بدین منظور، ابتدا به شناسایی فاکتورهای اصلی تاثیرگذار بر تحلیل کسب و کار سازمان از منظر رقابت و بازاریابی می پردازیم. سپس شناسایی مکانیزم فعالیت و چارچوب های کاربردی Big Data را بررسی خواهیم کرد. پژوهش حاضر با بیان ویژگی ها و خصوصیات منحصر به فرد ابزار داده های عظیم یا همان Big Data توصیفی تلقی گردیده و با در نظر گرفتن این امر که نتایج حاصل از آن در تمامی صنعت ها و فعالیت های اقتصادی در هر زمینه ای قابل بهره برداری و استفاده می باشد کاربردی قلمداد می گردد. به منظور انجام پژوهش پرسشنامه ای بین جامعه آماری توزیع گردید و با استفاده از نرم افزار SPSS مورد تجزیه و تحلیل قرار گرفت. یافته های حاصل از پژوهش نشان داد که فرضیه اول مبنی بر بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد، به علت آنکه مقدار سطح معناداری کمتر از ۰/۰۵ است؛ بنابراین بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

کلمات کلیدی: big data-تکنولوژی-فرآیند های بازار یابی-بهره برداری

**فصل اول: کلیات پژوهش** ..... ۱

۱-۱ مساله اصلی تحقیق : ..... ۲

۱-۲ تشریح و بیان موضوع : ..... ۲

۱-۳ ضرورت انجام تحقیق : ..... ۳

۱-۴ اسابقه تحقیقات و مطالعات انجام گرفته : ..... **Error! Bookmark not defined.**

۱-۵ اهداف اساسی از انجام تحقیق : ..... ۴

۱-۶ نتایج مورد انتظار پس از انجام این تحقیق : ..... ۴

۱-۷ روش انجام پژوهش : ..... ۵

۱-۷-۱ روش تحقیق : ..... ۵

۱-۷-۲ روشهای گردآوری اطلاعات : ..... ۵

۱-۸ اقلمرو تحقیق : ..... ۵

۱-۹ جامعه آماری : (N) ..... ۵

۱-۱۰ نمونه آماری : (n) ..... ۵

۱-۱۱ روش یا روشهای نمونه گیری : ..... ۵

۱-۱۲ روشهای مورد نظر برای تجزیه و تحلیل اطلاعات و آزمون فرضیه ها : ..... ۶

۱-۱۳ تعریف واژه ها و اصطلاحات تخصصی طرح : ..... ۶

**فصل دوم: مبانی نظری و پیشینه تحقیق** ..... ۷

۲-۱ رقابت و بازاریابی ؛ مؤلفه های مهم تحلیل کسب و کار ..... ۸



۹-۱-۲-۴	برنامه‌ریزی و پایش تحلیل کسب و کار	۹
۹-۱-۲-۵	استخراج نیازمندی‌ها	۹
۹-۱-۲-۶	مدیریت نیازمندی‌ها و ارتباطات	۱۰
۹-۱-۲-۹	اعتبارسنجی و ارزیابی راه‌حل	۱۰
۹-۱-۲-۱۲	جایگاه تحلیل کسب و کار در سازمان‌ها	۱۱
۲-۲	big data و شاخصه های آن	۱۷
۲-۲-۱	پیشینه پژوهش راجع به کلان داده‌ها	۱۹
۲-۲-۲	ارزش کلان داده ها :	۲۵
۲-۲-۴	چالش های کلان داده ها	۲۹
۲-۲-۵	نمایش داده	۳۵
۲-۲-۶	فشاردهی و کاهش افزونگی داده‌ها	۳۶
۲-۲-۷	مدیریت چرخه ی زندگی داده ها	۳۶
۲-۲-۷	مکانسیم تحلیلی	۳۶
۲-۲-۸	محرمانگی داده ها	۳۷
۲-۲-۹	مدیریت انرژی	۳۷
۲-۲-۱۰	توسعه پذیری و مقیاس پذیری	۳۷
۲-۲-۱۱	همکاری	۳۷
۲-۲-۱۲	پیدایش ارزش نظام مند کسب و کار	۳۸
۲-۲-۱۱	کشف دانش از داده های حجیم (KDD)	۳۹
۲-۳-۱	تأثیر big data بر فرآیند بهبود تحلیل کسب و کار	۴۱

## فصل سوم: روش تحقیق ..... ۶۵

۳-۱ انواع و ماهیت تحقیق ..... ۶۶

۳-۲ روش گردآوری اطلاعات ..... ۶۷

۳-۳ روش کتابخانه ای (توصیفی - پیمایشی) ..... ۶۸

۳-۴ روش مطالعه میدانی ..... ۶۸

۳-۵ ابزار گردآوری اطلاعات ..... ۶۸

۳-۶ پرسشنامه ..... ۶۹

۳-۹ روایی پرسشنامه ..... ۶۹

۳-۱۰ پایایی پرسشنامه (قابلیت اعتماد) ..... ۶۹

۳-۱۲ جامعه آماری ..... ۷۰

۳-۱۳ روش نمونه گیری ..... ۷۰

۳-۱۴ روش تعیین حجم نمونه ..... ۷۱

۳-۱۵ تجزیه و تحلیل اطلاعات ..... ۷۲

۳-۱۶ نحوه تفسیر ضریب همبستگی ..... ۷۲

## فصل چهارم: تجزیه و تحلیل داده ها ..... ۷۴

۴-۱ مقدمه ..... ۷۵

۴-۲ تحلیل های آماری توصیفی ..... ۷۵

۴-۲-۱ توزیع فراوانی بر حسب جنس ..... ۷۶

۴-۲-۲ توزیع فراوانی بر حسب سن ..... ۷۷

۴-۲-۳ توزیع فراوانی بر حسب تحصیلات ..... ۷۹

۴-۲-۵	توزیع فراوانی بر حسب سابقه خدمت	۸۰
۴-۳-۴	توصیف متغیرهای تحقیق	۸۱
۴-۳-۱	یکپارچگی	۸۱
۴-۳-۲	جذب مشتریان	۸۲
۴-۳-۳	تامین نیازها و خواسته های مشتریان	۸۳
۴-۳-۴	دسترس پذیری	۸۴
۴-۳-۵	مدیریت پایگاه داده	۸۵
۴-۳-۶	کاربرد Big Data	۸۶
۴-۴	تجزیه و تحلیل استنباطی داده ها	۸۷
۴-۵	بررسی طبیعی بودن توزیع داده ها	۸۸
۴-۵	جمع بندی فصل	۹۹
<b>فصل پنجم: بحث، نتیجه گیری و پیشنهادات</b>		
۵-۱	مقدمه	۱۰۱
۵-۳	بحث و نتیجه گیری	۱۰۱
۵-۴	محدودیت های پژوهش	۱۰۳
۵-۵	پیشنهادات	۱۰۳
۵-۵-۱	پیشنهادات کاربردی	۱۰۴
۵-۵-۲	پیشنهادات آتی	۱۰۴

## فهرست جداول

جدول ۴-۱ توزیع فراوانی بر حسب جنس	۷۶
نمودار ۴-۱ نمودار توزیع فراوانی بر حسب جنس	۷۷
جدول ۴-۲ توزیع فراوانی بر حسب سن	۷۷
نمودار ۴-۲ نمودار توزیع فراوانی بر حسب سن	۷۸
جدول ۴-۳ توزیع فراوانی بر حسب تحصیلات	۷۹
نمودار ۴-۳ نمودار توزیع فراوانی بر حسب تحصیلات	۸۰
جدول ۴-۵ توزیع فراوانی بر حسب سابقه خدمت	۸۰
نمودار ۴-۵ نمودار توزیع فراوانی بر حسب سابقه خدمت	۸۱
جدول ۴-۶ توصیف متغیر یکپارچگی	۸۲
نمودار ۴-۶ هیستوگرام متغیر یکپارچگی	۸۲
جدول ۴-۷ توصیف متغیر جذب مشتریان	۸۲
نمودار ۴-۷ هیستوگرام متغیر جذب مشتریان	۸۳
جدول ۴-۸ توصیف متغیر تامین نیازها و خواسته های مشتریان	۸۳
نمودار ۴-۸ هیستوگرام متغیر تامین نیازها و خواسته های مشتریان	۸۴
جدول ۴-۹ توصیف متغیر تامین دسترس پذیری	۸۴
نمودار ۴-۹ هیستوگرام متغیر دسترس پذیری	۸۵
جدول ۴-۱۰ توصیف متغیر مدیریت پایگاه داده	۸۶
نمودار ۴-۱۰ هیستوگرام متغیر مدیریت پایگاه داده	۸۶
جدول ۴-۱۱ توصیف متغیر کاربرد Big Data	۸۷
نمودار ۴-۱۱ هیستوگرام متغیر کاربرد Big Data	۸۷
جدول ۴-۱۲ تست نرمالیتت متغیرها	۸۸
جدول ۴-۱۳ ضریب همبستگی پیرسون فرضیه اول	۸۹

جدول ۴-۱۴	آزمون F (آنالیز واریانس) جهت معنادار بودن رگرسیون	۹۰
جدول ۴-۱۵	ضرایب تاثیر رگرسیون	۹۱
جدول ۴-۱۷	آزمون F (آنالیز واریانس) جهت معنادار بودن رگرسیون	۹۲
جدول ۴-۱۸	ضرایب تاثیر رگرسیون	۹۳
جدول ۴-۱۹	ضریب همبستگی پیرسون فرضیه سوم	۹۳
جدول ۴-۲۰	آزمون F (آنالیز واریانس) جهت معنادار بودن رگرسیون	۹۴
جدول ۴-۲۱	ضرایب تاثیر رگرسیون	۹۴
جدول ۴-۲۲	ضریب همبستگی پیرسون فرضیه چهارم	۹۵
جدول ۴-۲۳	آزمون F (آنالیز واریانس) جهت معنادار بودن رگرسیون	۹۶
جدول ۴-۲۴	ضرایب تاثیر رگرسیون	۹۶
جدول ۴-۲۵	ضریب همبستگی پیرسون فرضیه پنجم	۹۷
جدول ۴-۲۶	آزمون F (آنالیز واریانس) جهت معنادار بودن رگرسیون	۹۸
جدول ۴-۲۷	ضرایب تاثیر رگرسیون	۹۸
جدول ۵-۱	خلاصه نتایج	۱۰۱

## فصل اول: کلیات پژوهش

## ۱-۱ مساله اصلی تحقیق :

در دنیای پیچیده امروزی که مدام در حال تغییر و تحول حوزه های گوناگون فناوری و تکنولوژی می باشد ایجاد ظرفیت ها و توانایی های لازم جهت حفظ فضای کسب و کار و رقابت با سایر رقبای به امری کاملاً ضروری برای بقا تبدیل گردیده است. چرا که قابلیت پاسخ گویی به نیازهای متنوع و گوناگون مشتریان؛ نیازمند درک صحیح و شناخت این نیازها و انجام سریع اقدامات لازم با هدف عقب نماندن از فرصت های بازار می باشد.

با در نظر گرفتن شرایط حاکم بر بازارهای امروزی کاملاً می توان متوجه شد که دریافت و شناخت نیازهای دائماً در حال تغییر مشتریان و مردم مهم ترین فعالیت و اقدامات شرکت های فعال بازار را شکل می دهد چرا که هر کدام از این شرکت ها تنها در صورتی می توانند به حیات خود ادامه دهند که خود را با این تقاضا ها هماهنگ ساخته و گوی رقابت را از سایر رقبا برابیند.

سازمان ها و شرکت ها با به کارگیری و پیاده سازی فناوری اطلاعات (IT) توانایی انجام بهتر و ساده تر وظایف خود را افزایش می دهند و از این طریق قادرند روش کار خویش را دگرگون سازند. مزایایی که IT در سازمان ها ایجاد می کند از جمله در صرفه جویی هزینه ها، جلوگیری از خطاهای انسانی، افزایش بهبود کارایی و اثر بخشی سازمانی بسیار قابل تامل می باشد. به همین خاطر، امروزه سرانه هزینه IT به ازای یک نیروی انسانی به عنوان یکی از شاخص های توسعه ملی کشورها مطرح می شود (کمپ، ۲۰۱۴).

Big data اصطلاحی است برای مجموعه داده های حجیم که بزرگ، متنوع، با ساختار پیچیده و با دشواری هایی برای ذخیره سازی، تحلیل و تصویرسازی (نمایش)، پردازش های بیشتر یا نتایج می باشد. پروسه تحقیق بر روی داده های حجیم جهت آشکارسازی الگوهای مخفی و راز همبستگی ها، تجزیه و تحلیل big data نامیده میشود. این اطلاعات مفید برای سازمان ها و شرکت ها در جهت کسب بینش غنی تر و عمیق تر و موفقیت در رقابت کمک میکند. به همین دلیل اجراهای big data نیاز دارند تا در صورت امکان، تحلیل شوند و به طور دقیق اجرا شوند. این پژوهش به بررسی کاربرد BigData در تجزیه و تحلیل های مرتبط به کسب و کار شرکت ها در محیط رقابتی می پردازد (ایناو و لوین، ۲۰۱۳).

## ۲-۱ تشریح و بیان موضوع :

با سپری شدن سال های متمادی در صنعت IT و رشد تجارب گوناگون در حوزه جمع آوری، ذخیره سازی و بازیابی اطلاعات، موضوع معنا بخشیدن به داده ها و آسان کردن فرآیند تصمیم سازی، مرکز توجه کارشناسان فناوری اطلاعات و متخصصان علم مدیریت و کسب و کار قرار گرفته است. می دانیم که تصمیم سازی به معنای ساختن و پیشنهاد انواع تصمیماتی است که در شرایط خاص می توان اتخاذ کرد. بنابراین داده های عظیم یا همان Big Data قابلیت آن را دارند که با استفاده از تکنیک های پیشرفته ارزیابی و تحلیل داده ها، اطلاعاتی بسیار با ارزش و گرانبهایی را در اختیار شرکت ها قرار دهند. این تکنیک های پیشرفته که شامل طیف وسیعی از نرم افزار های کاربردی، مدل های کسب و کار و الگوهای مختلف می باشد به سازمان ها کمک می کند تا داده های پراکنده موجود را به اطلاعاتی کارگشا، قابل تبادل (به اشتراک گذاشتن میان مدیران) و قابل ذخیره سازی برای تصمیمات بعدی تبدیل کنند (مانیکا و همکاران، ۲۰۱۱).

امروز هر سازمانی که خواستار توسعه بازاریابی و فروش محصولات و خدمات خود در هر محدوده ای باشد، باید شناخت بازار را که مستلزم درک جامعی از نیازمندی های مشتریان و فضای عمومی عرضه و تقاضای تجارت در منطقه فعالیتش است را سرلوحه توسعه کسب و کار خود قرار دهد و به این ترتیب، فروش محصولات و خدمات شرکت مربوطه افزایش یافته و در نهایت اعتبار سازمان ارتقا می یابد. درواقع بازاریابی فرزند اقتصاد رقابتی است.

در دورانی که سرعت اتصال به اینترنت و حجم دستگاه های متصل به آن کم بود تحلیل سازمانی بیش تر بر مبنای ادراک مدیران و به صورت شهودی انجام می شد. اما به مرور بارشده انتظارات سازمان ها و مشتریان نسبت به دریافت خدمات سریع، تحلیل بلا درنگ به عنوان ابزاری ارزشمند معرفی گردید که قادر است به کسب و کار های گوناگون جهت اتخاذ تصمیمات هوشمندانه و همسو با نیازهای مشتریان به منظور صرفه جویی در وقت و هزینه، کمک شایانی نماید. تحلیل بلا درنگ در کنار رشد و نوآوری های ایجاد شده در صنعت تراشه سازی و سایر فناوری ها قادر به مرتب نمودن اطلاعات حجیم سازمان ها در عرض یک میلیونیم ثانیه بودند، ابداع گردید و دارای محبوبیتی فراگیر شد. در مقابل کارمندان سازمان ها می توانند براساس نوع انتخاب خود، از این فناوری به صورت لجستیک یا استراتژیک استفاده نمایند (کاسترو و همکاران، ۲۰۰۷).

### ۳-۱ ضرورت انجام تحقیق :

عبارت Big Data مدت ها است که برای اشاره به حجم های عظیمی از داده ها که توسط سازمان های بزرگی مانند گوگل یا ناسا ذخیره و تحلیل می شوند مورد استفاده قرار گیرد. اما به تازگی، این عبارت بیشتر برای اشاره به مجموعه های داده ای بزرگی استفاده می شود که به قدری بزرگ و حجیم هستند که با ابزارهای مدیریتی و پایگاه های داده سنتی و معمولی قابل مدیریت نیستند. مشکلات اصلی در کار با این نوع داده ها مربوط به برداشت و جمع آوری، ذخیره سازی، جست و جو، اشتراک گذاری، تحلیل و نمایش آن ها است. این مبحث، به این دلیل هر روز جذابیت و مقبولیت بیشتری پیدا می کند که با استفاده از تحلیل حجم های بیشتری از داده ها، می توان تحلیل های بهتر و پیشرفته تری را برای مقاصد مختلف، از جمله مقاصد تجاری، پزشکی و امنیتی، انجام داد و نتایج مناسب تری را دریافت کرد. بیشتر تحلیل های مورد نیاز در پردازش داده های عظیم، توسط دانشمندان در علوم مانند هواشناسی، ژنتیک، کاننومیک (علوم مرتبط با نگاشت سیستم عصبی)، شبیه سازی های پیچیده فیزیک، تحقیقات زیست شناسی و محیطی، جست و جوی اینترنت، تحلیل های اقتصادی و مالی و تجاری مورد استفاده قرار می گیرد (لوی و ویلنسکای، ۲۰۱۱).

حجم داده های ذخیره شده در مجموعه های داده ای Big Data، شبکه های حسگر بی سیم و غیره با سرعت خیره کننده ای در حال افزایش است، به طوری که در هر روز، ۲/۵ کوادریلیارد بایت (هر کوادریلیارد برابر ۱۰ به توان ۲۷ است) داده در حال تولید است. نکته جالب توجه در این زمینه آن است که ۹۰ درصد داده هایی که اکنون در اختیار ما است، تنها در ۲ سال اخیر تولید شده است!

یکی از مهم ترین مسائل مرتبط با داده های عظیم، مشکل بودن کار با آن ها به وسیله پایگاه های داده ای رابطه ای و بسته های نرم افزاری تصویر نگاری داده ها و نرم افزارهای آماری رومیزی است. چرا که این داده ها، برای پردازش شدن در یک زمان معقول به نرم افزارهای به شدت موازی شده با قابلیت اجرا روی ده ها، صد ها یا هزاران سرور نیاز دارند. البته مفهوم Big Data برای شرکت ها و سازمان های مختلف تعابیر متفاوتی دارد و هر کدام، بسته به کاربرد و نیازمندی



هایی که دارند، در حجمی خاص و با شرایطی خاص به روش های جدیدی برای آسان کردن کار با این نوع داده ها روی می آورند. از این رو است که Big Data برای بعضی سازمان ها، تنها صد ها گیگابایت حجم دارد در حالی که برای برخی، ده ها و صد ها ترابایت یا حتی مضاربی از اگزابایت وزتابایت از انواع داده های عظیم محسوب می شوند. یکی از بهترین تعابیری که در زمینه توصیف و تبیین Big Data و چالش های پیش روی آن به کار رفته است، ایده دوگ لنی (Doug Laney) در گزارش سال ۲۰۰۱ موسسه META group (اکنون گارتنر) بود که در آن عنوان شده است داده ها در سه بعد مختلف در حال رشد هستند (یک و مستو، ۲۰۰۸).

این سه بعد عبارتند از حجم، سرعت و تنوع داده ها که روز به روز نرخ رشد آن ها با سرعتی باور نکردنی افزایش می یابد. به همین دلیل، توصیف Big Data تنها با حجم های عظیم و نحوه مدیریت آن ها کار درستی نیست و باید دیگر جنبه های این مفهوم مهم و کلیدی را نیز در نظر گرفت.

بر این اساس، با توجه به افزایش روز افزون استفاده از تجهیزات تولید با جمع آوری داده ها و همچنین روی آوردن تعداد بیشتری از شرکت ها و افراد به شکل های جدیدی از زندگی دیجیتالی، اهمیت مفهوم Big Data و نحوه برنامه ریزی و تعیین راهبرد های مناسب برای بهره برداری صحیح از آن، دو چندان شده و نیاز به توسعه ابزارها و امکانات مناسب برای مدیریت بهتر آن ها بیش از پیش مشخص می شود (بیکر، ۲۰۱۲).

برنامه های بازاریابی، به تحلیل اطلاعات کلان داده ها درباره ی چگونگی جذب بیشتر مشتریان و ایجاد حس وفاداری و برقراری ارتباط پایدار با آن ها، نیاز دارند. ارتباطی که بتواند با پایداری در بلندمدت، نیازهای مشتریان را در مقابل رقبا، مرتفع کند. یافتن روش هایی برای عدم رغبت مشتریان به امتحان کردن یا انتخاب محصول و خدمات رقیب، استراتژی ای بسیار مهم تر از تبدیل مشتریان رقیب به مصرف کنندگان کالاها یا خدمات خود است. پاسخ سوالات اساسی و یافتن راه حل های مطلوب، نیازمند استفاده ی متفاوت از کلان داده ها است. لازم است به جای یافتن راه حل هایی برای هدف گیری مشتریان بالقوه، به دنبال استفاده از کلان داده ها جهت کسب صفات ارزشمند در محصولات و خدمات بود.

#### ۴-۱ اهداف اساسی از انجام تحقیق :

هدف این پژوهش تحلیل کاربرد Big Data بر فرآیند های بازاریابی و رقابت می باشد. به منظور نیل به این هدف، اهداف فرعی زیر تدوین گردیدند:

- شناسایی فاکتورهای اصلی تاثیرگذار بر تحلیل کسب و کار سازمان از منظر رقابت و بازاریابی
- شناسایی مکانیزم فعالیت و چارچوب های کاربردی Big Data

#### ۵-۱ نتایج مورد انتظار پس از انجام این تحقیق :

۱. شناخت کاربردهای Big Data در بازاریابی
۲. تاثیر آن در فرآیند های بازاریابی بهتر و موثرتر.

## ۶-۱ روش انجام پژوهش :

### ۶-۱-۱ روش تحقیق :

پژوهش حاضر با بیان ویژگی ها و خصوصیات منحصر به فرد ابزار داده های عظیم یا همان Big Data، از نظر هدف، کاربردیو از نظر جمع‌آوری داده، توصیفی از نوع پیمایشی تلقی گردیده و با در نظر گرفتن این امر که نتایج حاصل از آن در تمامی صنعت ها و فعالیت های اقتصادی در هر زمینه ای قابل بهره برداری و استفاده می باشد کاربردی قلمداد میگردد . ( حافظ نیا ۱۳۸۹ )

### ۶-۲ روشهای گردآوری اطلاعات :

شیوه جمع و گردآوری اطلاعات نیز با بهره گیری از روش کتابخانه ای و استفاده از منابع ، کتب و مقالات علمی و پرسشنامه می باشد. در پژوهش مورد نظر سعی بر آن می گردد که با ارائه تصویری جامع و دقیق از ابزار big data به عنوان عاملی جهت توفیق و رونق کسب و کار که شامل ماهیت توصیفی آن نیز است به هر چه کاربردی تر بودن نتایج حاصل از آن همت گمارده شود.

### ۶-۷ قلمرو تحقیق :

- قلمرو موضوعی: کاربرد ابزار فناوری اطلاعات Big Data در بازاریابی
- قلمرو مکانی: در شرکت کسب و کار توسعه منابع انرژی توان
- قلمرو زمانی: نیمه اول سال ۱۳۹۶

### ۸-۱ جامعه آماری : (N)

جامعه آماری این پژوهش را مدیران و کارشناسان بازاریابی و کارمندان شرکت توسعه منابع انرژی توان بوده که تعداد آنها ۶۰ نفر به عنوان جامعه آماری می باشند.

### ۹-۱ نمونه آماری : (n)

کارمند شرکت توسعه منابع انرژی توان

### ۱۰-۱ روش یا روشهای نمونه گیری :

نمونه گیری تصادفی

## ۱۱- روشهای مورد نظر برای تجزیه و تحلیل اطلاعات و آزمون فرضیه ها :

به منظور توصیف ویژگی های نمونه، ابتدا داده های جمع آوری شده با استفاده از شاخص های آمار توصیفی خلاصه و طبقه بندی می شود، و سپس با استفاده از شاخص های آمار استنباطی به تایید یا رد فرضیات پرداخته می شود. به منظور اجرای شیوه پیمایشی، بر اساس مطالعات صورت گرفته، گویه های ابزار پرسشنامه تدوین شد و پس از بررسی اساتید، پرسشنامه نهایی و تصحیح شد. پس از سنجش اعتبار پرسشنامه، از میان جامعه آماری تحقیق، با استفاده از روش های تعیین حجم نمونه، حجم نمونه آماری تحقیق حاضر محاسبه گردید. سپس با استفاده از روش نمونه گیری به شیوه تصادفی اقدام به توزیع پرسشنامه در میان جامعه آماری شده است. پس از گردآوری پرسشنامه های تکمیل شده، اقدام به نمره دهی پاسخ ایشان خواهد شد. پس از نمره دهی هر گویه، به منظور تجزیه و تحلیل داده های آماری از نرم افزار Spss و آزمون های آماری مناسب استفاده خواهد شد. سپس در مرحله تدوین فصل چهارم اقدام به توضیح آزمون ها و نتایج حاصله در قالب جداول و نمودارهایی شد. در نهایت بر اساس نتایج مطالعه حاضر و نیز مطالعات دیگران، پیشنهاداتی ارائه شد.

## ۱۲- تعریف واژه ها و اصطلاحات تخصصی طرح :

**بازاریابی:** به عنوان فرایندی مدیریتی... اجتماعی تعریف می شود که بوسیله آن افراد و گروه ها از طریق تولید و مبادله کالا با یکدیگر، به امر تأمین نیازها و خواسته های خود اقدام می کنند.

**مدیریت بازاریابی:** تلاش آگاهانه ای است برای به دست آوردن نتایج مثبت در مبادله با بازارهای هدف.

**Big data یا کلان داده ها:** کلان داده دارایی های داده ای اند بسیار انبوه، پرشتاب و گوناگون که نیاز به روش های پردازشی تازه ای دارند تا تصمیم گیری، بینش تازه و بهینگی پردازش پیشرفته را فراهم آورند.

**رقابت:** در علم اقتصاد بین فروشندگان در تلاش برای دستیابی به اهدافی چون سود فزاینده، سهم بازار، و حجم فروش از طریق تغییر دادن عناصر ترکیب بازاریابی انجام می گیرد.

## فصل دوم: مبانی نظری و پیشینه تحقیق

## ۱-۲ رقابت و بازاریابی؛ مؤلفه های مهم تحلیل کسب و کار

امروزه جهان بیش از هر زمان دیگری پیچیده شده و این پیچیدگی ها تنها به خاطر پیشرفت علوم و فناوری نیست. بلکه بخشی از آن به دلیل افزایش تعداد پارامترهای تأثیرگذار بر یک مسئله می باشد. همین امر خود باعث ایجاد تخصص های حل مسئله، تصمیم گیری و تصمیم سازی شده است. در این بین دنیای کسب و کار هم تأثیر زیادی از این مسائل پذیرفته و علاوه بر همه این موارد رشد و توسعه روزافزون فناوری اطلاعات و جایگاه آن در سازمان ها و کسب و کارها نیز خود باعث پیچیدگی مضاعف موضوع گردیده است.

اصطلاح کسب و کار معمولاً این روزها معادل کلمه بیزنس در زبان انگلیسی به کار می رود. یکی از معانی بیزنس، کسب و کار یا تجارت است. این کلمه معانی و مفهوم گسترده تری، به ویژه در ادبیات سیستمی، دارد. می توان این طور گفت که فعالیت یا فعالیت های یک سازمان یا بنگاه اقتصادی و حتی غیر اقتصادی، مجموعاً کسب و کار یا بیزنس آن را تشکیل می دهند. بنابراین در اینجا منظور ما از کسب و کار معنای عام و اولیه آن یعنی خرید و فروش و فعالیت محض اقتصادی نیست (لینگس، ۲۰۱۴).

یک سازمان یا بنگاه برای انجام فعالیت های خود، چه اقتصادی و چه غیر اقتصادی، نیاز به یک ساختار و نظم مشخص دارد. بدیهی است بدون داشتن ساختار و نظم بنگاه قادر به انجام فعالیت های خود نخواهد بود. یک مغازه خشکبار فروشی کوچک را در نظر بگیرید. صاحب این مغازه برای اداره موفق کسب و کار خود، اولین اقدامی که انجام می دهد اختصاص یک دفتر کوچک حساب و کتاب برای ثبت اطلاعات حساب مشتریان خود است. دومین گام ثبت و نگهداری خریدها و فروش های مغازه است. سومین اقدام شناسایی مشتریان و ارزیابی اعتبار آنها ولو به صورت ذهنی و نانوشته می باشد. گرچه امروزه نرم افزارهای کاربردی کوچک همه این موارد را به راحتی مدیریت می کنند، ولی همین مثال ساده مشخص می نماید که حتی یک مغازه کوچک هم برای خودش ساختار و نظم مشخصی دارد که البته همیشه مکتوب و نوشته شده نیست.

ساختار و نظم که در بالا به آن اشاره شد می تواند مواردی همچون: مدیریت کسب و کار، مدیریت بازار، مدیریت مشتریان، مدیریت نام تجاری، مدیریت تامین کنندگان و ... باشد. مجموعه این موارد است که مفهومی با نام کسب و کار را پدید می آورد.

تحلیل کسب و کار مجموعه ای از وظایف و تکنیک هایی است که برای برقراری ارتباط میان ذینفعان به منظور فهم ساختارها، سیاست ها و عملیات یک سازمان مورد استفاده قرار گرفته، و راه حل هایی را برای توانمندسازی سازمان در راستای دستیابی به اهدافش پیشنهاد می دهد. در واقع تحلیل کسب و کار درک این موضوع است که چگونه یک سازمان برای دستیابی توأم با موفقیت به اهدافش عمل نماید. و همچنین به چه صورت قابلیت ها یا توانمندی های مورد نیاز جهت ارائه محصولات و خدمات به ذینفعان برون سازمانی نظیر مشتریان را تعریف کند. علاوه بر این ها در تحلیل کسب و کار اهداف

کلان سازمانی و اهداف میان مدت، طبقه ارتباط این اهداف با هم، هم‌راستا بودن یا نبودن اهداف میان مدت و کلان، محقق‌سازی آن‌ها و رشته فعالیت‌های مورد نیازی که باید انجام شوند نیز تعریف می‌گردند.

تحلیل کسب و کار ممکن است با اهداف متفاوتی انجام شود. برخی مواقع تحلیل کسب و کار به منظور شناخت وضعیت موجود یا فعلی سازمان صورت می‌پذیرد. شناخت وضعیت موجود می‌تواند به عنوان بستری برای شناسایی نیازهای آتی کسب و کار عمل نماید. گرچه بیشتر مواقع تحلیل کسب و کار برای تعریف و صحت سنجی راه‌حل‌هایی که نیازها، اهداف میان مدت و کلان سازمان را پوشش می‌دهند، مورد استفاده قرار می‌گیرد.

کاری که تحلیلگر کسب و کار انجام می‌دهد، دقیقاً حلاجی کردن اطلاعاتی است که از منابع مختلف نظیر مشتریان، کارکنان سازمان، متخصصان فناوری اطلاعات و مدیران جمع‌آوری شده است. در بیشتر موارد تحلیلگر کسب و کار نقش یک روان‌ساز یا تسهیل‌کننده ارتباطات بین واحدهای سازمانی را ایفا می‌نماید. به ویژه این نقش در هم‌راستا کردن نیازهای کسب و کار با قابلیت‌های راه‌حل‌های مبتنی بر فناوری اطلاعات پررنگ تر شده و به نوعی تحلیلگر در اینجا به عنوان یک مترجم یا مفسری که نیازهای کسب و کار را به زبان قابل فهم متخصصان IT ترجمه می‌نماید، حضور پیدا می‌کند.

#### ۴-۱-۲ برنامه‌ریزی و پایش تحلیل کسب و کار

این حوزه دانشی از دو قسمت مهم و البته مکمل، برنامه‌ریزی و پایش تشکیل شده است. حوزه BAP&M به دنبال ایجاد یک بستر مناسب برای تعریف، برنامه‌ریزی و اتمام یک پروژه تحلیل کسب و کار به صورت موفقیت آمیز می‌باشد. بخش برنامه‌ریزی در این حوزه بر روی آماده‌سازی برای اجرای مؤثر یک پروژه تحلیل کسب و کار تمرکز می‌نماید. ایجاد یک درک عمیق و مناسب از ماهیت پروژه، نیازمندی‌های کسب و کار و اهداف میان مدت امکان انتخاب یک رویکرد مناسب در مسیری درست و سپس توسعه یک راه‌حل برای نیازهای شناسایی شده را فراهم می‌آورد. این بُعد شامل تحلیل ذینفعان، نقش‌ها و مسئولیت‌های آن‌ها در پروژه تحلیل، نیازهای آن‌ها برای اطلاعات درباره فعالیت‌های تحلیل کسب و کار، و نیز سطح جزئیات مورد نیاز و تعداد ارتباطات برای کسب اطلاعاتی مشخص می‌باشد.

بخش پایش در این حوزه برای حصول اطمینان از این است که آیا اجرای فعالیت‌ها مطابق با برنامه‌های توافق شده بوده‌اند یا خیر؟ و همچنین این که فرآورده‌های تولید شده همان مواردی بوده‌اند که در برنامه‌های اجرایی بر روی آن‌ها توافق شده است؟ در این حوزه معیارهای مورد نیاز برای پایش فعالیت‌های تحلیل کسب و کار نیز تعیین می‌گردند.

برخی از وظایف این حوزه دانش در خلال پیشرفت پروژه و همگام با آن ممکن است کامل شده یا دستخوش تغییر شوند. برخی از وظایف مشخص شده در این حوزه نیز با عنوان فعالیت‌هایی که قبل از آغاز پروژه تحلیل باید انجام شده باشند، مشخص می‌شوند (فیتزگراالد، ۲۰۱۲).

## ۲-۱-۵ استخراج نیازمندی‌ها

استخراج نیازمندی‌ها یکی از فعالیت‌های کلیدی تحلیل کسب و کار است. استخراج نیازمندی‌ها تحلیل گر کسب و کار را به این سو که ذینفعان چه نیاز یا انتظاری از راه‌حل جدید دارند، هدایت می‌نماید. در این حوزه دانشی به کمک تکنیک‌های مشخص و متداول نیازمندی‌های مطرح در سطوح مختلف استخراج می‌گردند. یکی از موضوعات مهم در این حوزه انتخاب تکنیک مناسب توسط تحلیل گر برای استخراج اطلاعات است. مسلماً بدون داشتن اطلاعات درست و مؤثر فعالیت‌های این حوزه ناقص صورت خواهد پذیرفت. بدیهی است که بدون داشتن اطلاعات درست یا مورد وثوق در سازمان پروژه ناموفق خواهد بود. دقت شود که انتخاب منابع درست اطلاعاتی، اطلاعات درست را در پی خواهد داشت. و اطلاعات درست هم توسعه نیازمندی‌های درست (یا صحیح) را به دنبال دارد.

استخراج اطلاعات معمولاً برای اولین بار، در ابتدای چرخه عمر پروژه و زود هنگام انجام می‌گیرد. اما در خلال پروژه همراه با تغییر نیازمندی‌ها یا نیازمندی‌های فراموش شده باز هم استخراج صورت خواهد گرفت (کارکنولد، ۲۰۱۰).

## ۲-۱-۶ مدیریت نیازمندی‌ها و ارتباطات

پر واضح است که در یک پروژه تحلیل کسب و کار طیفی از ذینفعان از کاربران نهایی گرفته تا مدیران ارشد حضور دارند. برای انجام موفق یک پروژه تحلیل کسب و کار ایجاد و مدیریت یک درک مشترک از نیازمندی‌ها برای مخاطبان ضرورت دارد. بر این اساس در حوزه مدیریت نیازمندی‌ها و ارتباطات اطمینان حاصل می‌گردد که همه ذینفعان به یک برداشت مشترک از محتوای راه‌حل رسیده‌اند. این حوزه مشخص می‌نماید که کدام مجموعه از نیازمندی‌ها با یک گروه خاص از ذینفعان مرتبط بوده و چگونه آن نیازمندی‌ها را باید در قالبی مناسب برای مخاطبان ارائه داد.

مدیریت نیازمندی‌ها با موضوعاتی نظیر مدیریت تغییرات، و تأثیرات ناشی از این تغییرات بر سایر نیازمندی‌های مرتبط سروکار دارد. همچنین بسته‌بندی نیازمندی‌ها به صورتی منطقی و مناسب و نگهداری از آن‌ها به منظور استفاده مجدد در آینده از موضوعاتی است که در اینجا به آن توجه می‌شود (پریسمن، ۲۰۱۲).

## ۲-۱-۹ اعتبارسنجی و ارزیابی راه‌حل

در حوزه دانش اعتبارسنجی و ارزیابی راه‌حل اطمینان حاصل می‌گردد که راه‌حل ارائه شده نیازهای کسب و کار را پوشش داده و امکان پیاده سازی موفق آن‌ها را فراهم می‌آورد. فعالیت‌هایی که در این حوزه انجام می‌شوند می‌توانند شامل مواردی همانند ارزیابی و اعتبارسنجی فرآیندهای کسب و کار، ساختارهای سازمانی، توافق نامه‌های برون سپاری، نرم افزارهای کاربردی و یا هر یک از دیگر اجزاء راه حل باشند.

تحلیل گر کسب و کار نقشی بسیار مهم در حصول اطمینان از فرآیندهای مرور، انتخاب و طراحی راه حل کسب و کار در مسیری را دارد که ارزش قابل ارائه به ذینفعان را به حداکثر برساند. به همین منظور باید فعالیت‌های این حوزه دانش توسط وی و تیم تحلیل صورت بگیرند.

## ۲-۱-۱۲ جایگاه تحلیل کسب و کار در سازمان‌ها

کسب و کار به مثابه یک سیستم و بالاتر از آن یک سیستم ارگانیکی می‌باشد. بنابراین برخوردی که با این سیستم می‌گردد، باید از جنس برخوردی باشد که با یک سیستم ارگانیکی می‌شود. پویایی و زنده بودن امروزه بیش از پیش در تار و پود کسب و کارها نفوذ کرده است. بنابراین فرمول‌های گذشته برای تحلیل کسب و کار دیگر کارساز نخواهند بود.

معمولاً زمانی که از تحلیل کسب و کار نام برده می‌شود، یک طیف به این صورت که در یک سر آن فعالیت‌های محض کسب و کار مثل فعالیت‌های مالی، بازرگانی، مهندسی و غیره قرار داشته و سر دیگر آن که فعالیت‌های محض فناوری اطلاعات مثل توسعه سیستم، معماری نرم افزار یا سازمان و غیره قرار می‌گیرند تصویر می‌شود. علاوه بر این اگر محیط کسب و کارها را بخواهیم از نظر میزان تأثیرپذیری عوامل داخلی و خارجی مورد بررسی قرار دهیم، به این نتیجه خواهیم رسید که تأثیرپذیری کسب و کارها از عوامل خارجی روز به روز در حال افزایش است. حالتی که در سالیان گذشته نمود نداشته و یا به این شدت نبوده است. موضوع دیگری که سالیان نه چندان زیادی در ذات کسب و کارها نفوذ کرده و عملاً به عنوان جزء جدایی‌ناپذیر آن‌ها قلمداد می‌گردد بحث فناوری اطلاعات و ابزارها و راه‌حل‌های مبتنی بر فناوری اطلاعات است (اسوینای و سوتار، ۲۰۰۱).

بر اساس مطالبی که تا اینجا بیان گردیدند، اگر در گذشته تنها به کسب و کار از یک بعد توجه می‌شده، اگر کسب و کارها تنها از زاویه عوامل داخلی بررسی می‌شده‌اند، اگر راه‌حل‌های فناوری اطلاعات اصلاً در کسب و کارها حضور نداشته و یا اگر هم حضور داشته در راستای اهداف کسب و کار و همگرا با آن دیده نمی‌شده، امروزه اوضاع دگرگون شده است و کسب و کارها در این وضعیت دیگرگونه نگاه و مَنشی جدید را طلب می‌کنند.

امروزه نقش تحلیل گر کسب و کار در سازمان‌ها یک نقش اساسی و کلیدی است. آن‌ها به عنوان حل‌کننده مسائل و مشکلات سازمان مطرح می‌باشند. زمانی بود که این افراد به عنوان نقشی در سر ابتدایی طیفی که در بالا عنوان گردید، قرار می‌گرفتند. مثلاً تحلیل گر صرف حوزه‌های مالی، بازرگانی و غیره. گرچه امروزه همچنان این پست‌ها در سازمان‌ها وجود دارد ولی شاید بتوان گفت طی این سال‌ها و با ورود و گسترش فناوری اطلاعات نوعی استحاله صورت پذیرفته است. بر اساس تحقیقی که محققان مؤسسه فورستر در سال ۲۰۰۸ انجام داده‌اند، امروزه نقش تحلیل‌گران کسب و کار همانند سایر نقش‌های فناوری محور، از یک تحلیل گر خاص یکی از حوزه‌های کسب و کار یا تحلیل گر محض مسائل فناورانه فراتر رفته است. این تحقیق بیان نموده است که طی این سال‌ها یک نقش جدید سازمانی با عنوان تحلیل گر فناوری کسب و کار شکل گرفته است که در وسط طیفی که گفته شد قرار می‌گیرد. این نقش هم از مسائل دامنه حوزه کسب و کار مطلع



است و هم از فناوری اطلاعات. بدین معنا که بیشتر حالتی بین‌وظیفه‌ای و بین‌دامنه‌ای داشته تا اینکه بر روی یک حوزه یا وظیفه مشخص از کسب و کار تمرکز نماید.

در بررسی دیگری که مجله کامپیوتر ورلد برای پیش‌بینی سال ۲۰۱۲ انجام داده بود، ۲۳٪ از سوال‌شوندگان پاسخ داده بودند که قصد دارند در ۱۲ ماه آینده افرادی با مهارت تحلیل کسب و کار را استخدام نمایند. این در حالی است که این عدد در تحقیق سال گذشته همین مجله ۱۳٪ بوده است.

در تحقیق مجله مانی متعلق به مؤسسه رسانه‌ای سی. ان. ان. تحلیلگر فناوری اطلاعات کسب و کار در سال ۲۰۱۱ به عنوان یازدهمین شغل از نظر میزان رشد در ۱۰ سال اخیر قرار داشته و رشدی معادل ۲۰٪ داشته است. این در حالی است که میزان رشد همین شغل در سال ۲۰۱۲ به عدد ۲۲/۱٪ رسیده است (لیو و همکاران، ۲۰۱۴).

مارک مک دونالد از تحلیل‌گران مؤسسه گارتنر معتقد است نگاه قدیمی به تحلیل‌گران کسب و کار به عنوان افرادی دون پایه در سازمان‌ها و کسانی که صرفاً به دنبال یادداشت برداری و مستند کردن جزئیات و طراحی BOM برای یک محصول بوده‌اند تغییر یافته و امروزه تحلیل‌گران کسب و کار با عنوان پرسنل ارشد و حل‌کننده‌های مسئله در سازمان‌ها مطرح می‌باشند. از نظر ایشان این تغییر بنا به دلیل سه چرخش زیر رخ داده است:

## ۲-۱-۱۳ تحلیل کسب و کار در توسعه سیستم‌های نرم‌افزاری

تحلیل کسب و کار معمولاً با هدف ارائه یک راه‌حل به منظور پاسخگویی به نیازهای سازمان انجام می‌پذیرد. این راه‌حل می‌تواند یک سیستم منطقی برای پیاده‌سازی در سازمان بوده یا یک راهکار مبتنی بر فناوری اطلاعات باشد. در این راستا اهمیت فناوری اطلاعات از یک سو و نقش آن به عنوان ابزاری برای محقق‌سازی منطق‌های مورد استفاده در کسب و کار از سویی دیگر باعث شده است تا خود به خود وقتی صحبت از راه‌حل می‌شود، یک راه‌حل مبتنی بر فناوری اطلاعات به ذهن متبادر گردد. گرچه کلیت این مسئله درست نیست، ولی بنا به دلایلی که گفته شد نقش این راه‌حل‌ها در سازمان و کسب و کارها بسیار مهم و حیاتی است. حال با توجه به اینکه نرم‌افزار به عنوان گل سرسبد راه کارهای مبتنی بر فناوری اطلاعات مطرح بوده و توسعه آن به شدت وابسته به فعالیت‌های بالادستی خود، مثل فعالیت‌های تحلیل کسب و کار، می‌باشد نقش و جایگاه این دو موضوع را به طور خلاصه مرور خواهیم کرد (لابرینیدیس و جاگادیش، ۲۰۱۲).

معمولاً مهندسان نرم‌افزار قبل از توسعه نرم‌افزار، کلیت سیستم را مهندسی می‌نمایند. فعالیتی که اصطلاحاً به آن مهندسی سیستم گفته می‌شود، بدین معناست که تمامی اجزاء سیستم شامل سخت‌افزار، نرم‌افزار، ذینفعان، پایگاه‌های داده، فرآیندها و غیره شناسایی، استخراج، تحلیل، مدل‌سازی، اعتبارسنجی و مدیریت می‌گردند.

به منظور پیاده‌سازی یک سیستم نرم‌افزاری نیاز است تا دامنه مسئله به دقت و با حساسیت تجزیه و تحلیل گشته و نهایتاً یک مدل منطقی قابل فهم برای تیم توسعه آماده نمود تا بتوان بر روی آن فناوری سوار نموده و به آن جان بخشید. جان بخشیدن در اینجا به معنای خارج کردن از حالت منطقی و عینیت دادن به مدل است. اگر بخواهیم ساده‌تر صحبت کنیم، می‌توان گفت تبدیل کردن مدل به یک برنامه نرم‌افزاری قابل اجرا در کامپیوتر.

سلسه مراتب مهندسی کردن سیستم، که خود حکایت از نوعی انتزاع نیز دارد، به صورت شکل (۲) قابل مدل‌سازی می‌باشد. شایان ذکر است قبل از انجام فعالیت‌های مرتبط با مهندسی نرم‌افزار باید واحد یا واحد های سازمانی مورد تحلیل قرار گیرند. همچنین بالاترین سطح سازمان مورد مطالعه و واحدهای آن مشاهده می‌گردند. این شکل به خوبی جایگاه تحلیل کسب و کار و مهندسی کردن آن تا توسعه سیستم نرم‌افزاری مورد نظر برای استقرار در این سازمان را نمایش می‌دهد.

قبل از ادامه این بحث لازم است تا به سرعت اتفاقی که چندسالی است در کسب و کارها افتاده است را مرور نماییم. زمانی محصولاتی در بازار صرفاً به اتکای ویژگی‌های فنی یا شاخص آن‌ها، بی‌رقیب و ممتاز بودند. رفته رفته این وضعیت تغییر حالت داد تا اینکه امروزه ممتاز بودن یک محصول به ارزشی که به واسطه آن نصیب مصرف‌کننده می‌نماید، مشخص می‌گردد. دیگر استفاده صرف از آخرین فناوری‌های روز یا گرانترین مواد اولیه و غیره امتیاز محسوب نمی‌شود بلکه خلق زنجیره ارزش برای مشتری یا فراتر از آن برای کسب و کار به واسطه استفاده از آن محصول است که مورد توجه قرار گرفته است. امروزه ما دیگر به ابزارها و خدمات تجارت الکترونیکی تنها به این دلیل که [amazon.com](http://amazon.com) را پیاده‌سازی کنیم، نیاز نداریم. بلکه می‌خواهیم فناوری اطلاعات به نیازهای کسب و کار پاسخ داده و راحتی تجارت و افزایش سود را به دنبال داشته باشد. شاید همه این‌ها موضوعاتی بودند که باعث شدند شرکت آمازون چند سال قبل تغییر راهبرد داده و از کتاب فروش بودن صرف دست کشیده و خود را به عنوان ارائه‌کننده خدمات تجارت الکترونیکی معرفی نماید. و در اقدامی تعجب برانگیز کتاب فروش های سنتی را نیز دعوت به استفاده از این گونه خدمات کرد.

همه این شواهد و قرائن بیان‌کننده یک چیز هستند و آن اینکه سازمان‌ها و شرکت‌های توسعه‌دهنده نرم‌افزار تنها با داشتن برنامه نویسان و طراحان پایگاه‌های داده و متخصصان امنیت نرم‌افزار خوب و حرفه‌ای، دیگر راه به جایی نخواهند برد. چاره کار استفاده از تحلیل کسب و کار و تحلیل‌گران کسب و کار برای ارائه راه‌حل‌های نرم‌افزاری موفق است. نرم‌افزارهایی که نیازهای ذینفعان خود را به خوبی پاسخ دهند.

علاوه بر مواردی که گفته شدند، امروزه نیازهای فناوری اطلاعات در سازمان‌ها نظیر راه‌حل‌های برنامه‌ریزی منابع سازمانی نه تنها در واحدهای متفاوت بلکه در کل سازمان گسترده و پخش شده‌اند. همچنین ورود فناوری‌های نرم افزاری جدیدتر همانند معماری سرویس‌گرا نیاز به درک عمیق از هر دو حوزه کسب و کار و فناوری اطلاعات را ضروری می‌سازد. طبیعتاً سازمانی نیز که در پی توسعه نرم افزارهایی بر پایه این گونه ابزارها و فناوری‌هاست، باید از متخصصانی استفاده نماید که به هر دو حوزه تسلط داشته باشند. یا اینکه با ادبیات دو حوزه آشنا بوده و بتوانند بین آن‌ها ارتباط برقرار نمایند. مسلماً با توضیحات داده شده این کار از عهده تحلیل‌گر کسب و کار فارغ از عنوان و سمت سازمانی وی، بر خواهد آمد.

می‌توان گفت که کسب و کارها امروزه برای موفقیت خود نیاز به استفاده از فناوریهای اطلاعاتی دارند. و فناوری اطلاعات نیز همانند هر فناوری و ابزار دیگری که در سازمان به کار گرفته می‌شود، باید به کسب و کار سازمان ارزش خاصی را ارائه نماید. برای این که بتوان راه‌حلی مبتنی بر فناوری اطلاعات برای کسب و کارها ارائه نمود که در عین فناوریانه بودنش ارزش مورد انتظار را نیز تحویل دهد، می‌بایست راه‌حل مورد نظر نیازمندی‌های کسب و کار و ذینفعان را پوشش دهد. و تنها رسیدن به این سطح از رضایت ذینفعان، تحلیل کسب و کار می‌باشد. امروزه استفاده از تحلیل کسب و کار با تعریف خاص خود و استاندارد که IIBA ارائه می‌دهد در کسب و کارها گریز ناپذیر شده است. سازمان‌ها باید به این سمت حرکت نمایند که هیچ راه‌حلی را بدون تحلیل دقیق شرایط و کسب و کارشان مورد پذیرش قرار ندهند. سازمان‌هایی که این هدف را دنبال می‌کنند می‌توانند از استاندارد BABOK و بهره‌گیری از آن در کسب و کارشان استفاده نمایند.

## \_\_بازاریابی

از دیدگاه فیلیپ کاتلر پدر علم بازاریابی نوین؛ بازاریابی به عنوان فرآیندی مدیریتی – اجتماعی تعریف می‌شود که بوسیله آن افراد، گروه‌ها و سازمان‌ها از طریق تولید و مبادله کالا(و خدمات) با یکدیگر، به تأمین نیازها و خواسته‌های خود اقدام می‌کنند. برای روشن شدن این تعریف باید در ابتدا نیاز، خواسته، تقاضا، کالا، مبادله، معامله و بازار را بررسی کرد:

نیاز یا need :

به نحوی احساس محرومیت انسان نسب به برخورداری از یک خدمت یا کالا را نیاز یا Need می‌گوییم.

خواسته یا want :

وقتی یک نیاز برای انسان برجسته شود تبدیل به خواسته می‌شود.

خواسته در علم اقتصاد و بازاریابی شکل ظاهری نیازهای ماست، که می تواند به صورت کالا یا خدمت مطرح شود، در واقع خواسته شکلی است که نیازهای انسان تحت تاثیر فرهنگ و شخصیت آنها به خود می گیرد، است.

خواسته بر اساس توانمندی تامین کننده یا تولید کننده ها در نحوه برطرف کردن نیاز تعریف می شود.

تقاضا یا demand :

در دنیای اقتصاد، تقاضا یعنی میل و خواست و توانایی یک فرد برای دریافت کالا یا خدمتی.

کالا یا good :

از دیدگاه علم اقتصاد، محصولات و مواد مختلفی که توسط تولیدکننده، به بازار یا Market عرضه شود و در برابر دریافت پول، یکی از نیازهای انسان را تأمین و رفع کند، کالا نام دارد.

نیاز، خواسته و تقاضا انسان را بسوی کالاهایی می کشاند که می تواند این نیازها و خواسته ها و تقاضاهایش را برآورده سازد. کالاهایی که در بازار عرضه می شوند، و به فروش می رسند تا توسط مصرف کنندگان خریداری شوند و به مصرف برسند، کالاهای مصرفی نام دارند و کالاهایی که تولیدکنندگان دیگری برای تولید کالاهای مختلف دیگر، آنها را خریداری کنند و مورد استفاده قرار دهند، کالای واسطه ای نامیده می شود. کالاهایی که نیازهای اولیه مصرف کنندگان را تأمین می کند، کالاهای ضروری و کالاهایی که برای تأمین نیازهای کم اهمیت تر مصرف می شود، کالاهای تجملی نامیده می شود.

## - مدیریت بازاریابی :

مدیریت بازاریابی فرآیند کاملی از برنامه ریزی ، ساماندهی ، تأمین نیروی انسانی ، آموزش و هدایت و رهبری مجموعه فعالیت های پژوهشی و بررسی در مورد بازاریابی ، نیازهای مشتریان ، وضعیت رقبا ، تولید و عرضه کالاهای دارای پتانسیل فروش و سعی در پیشبرد فروش آنها از طریق ایجاد تنوع و تغییرات لازم در زمان و مکان برای رسیدن به هدف حداکثر سود و تسلط به بازار از طریق به کار بستن بازاریابی است . فعالیت های مدیریت بازاریابی می بایستی در چارچوب برنامه های استراتژیک یک شرکت و با نگرش نظام گرا انجام گیرد .

مدیریت بازاریابی وظایف زیر را به عهده دارد :

۱. جمع آوری و بررسی اطلاعات به منظور تشخیص وضعیت و استعداد بازار و به طور کلی تعیین نیاز مصرف کنندگان مورد نظر . لازم به توضیح می باشد که مدیران بازاریابی می بایستی از طریق تقسیم بازار بتوانند مصرف کنندگان مورد نظر را شناسایی نمایند و با توجه به خصوصیات مصرف کنندگان ، کالا و خدمات مورد نیاز آنها را تأمین کنند .

۲. تعیین هدف های بازاریابی با توجه به نتایج اطلاعات به دست آمده و منابع شرکت مانند مشخص نمودن هدف های فروش و سهمیه بازار .
۳. تعیین آمیزه و یا ترکیب بازاریابی که عناصر عمده آن عبارتند از محصول ، قیمت ، کانال های توزیع و فعالیت های پیشبرد فروش ، که این عناصر به عنوان فعالیت های قابل کنترل مدیریت بازاریابی می باشند . مدیران بازاریابی میبایستی به این سوالات جواب دهند که چه محصولی با چه قیمتی تولید و در چه مکانی عرضه نمایند و با چه شیوه تبلیغاتی به اطلاع مصرف کنندگان برسانند. بدیهی است که این عوامل قابل کنترل مدیریت بازاریابی ، با توجه به شرایط محیطی ( سیاسی ، اقتصادی ، بازرگانی و فرهنگی ) که بر فعالیت های اقتصادی بنگاه تأثیر دارند و امکانات مؤسسه مشخص می گردند . نتیجتاً مدیریت بازاریابی با توجه به نیاز مصرف کنندگان و منابع سازمان می بایستی کالا یا خدمات مورد نیاز مصرف کنندگان را تهیه و عرضه نمایند.
۴. ساماندهی ؛ به منظور اینکه مدیریت بازاریابی بتواند به اهداف مورد نظر برسد می بایستی فعالیت های بازار یابی خود را سازماندهی نماید و نیروی انسانی و آموزش لازم را تأمین کند و ضمناً جهت پیشبرد فعالیت ها هماهنگی های لازم را داشته باشد .
۵. کنترل یا نظارت ؛ به منظور اینکه فعالیت ها با توجه به برنامه پیش برود ، مدیریت بازاریابی می بایستی در مقاطع زمانی مختلف عملکرد را با برنامه مورد بررسی و مقایسه قرار دهد و در جهت رفع مشکلات و تجدیدنظر در برنامه اقدام کند .
۶. تأمین خدمات بعد از فروش ؛ مثل تأمین لوازم و قطعات و ارائه خدمات تعمیراتی جهت محصولاتی که به فروش رفته اند . بدیهی است مصرف کنندگان در تصمیم گیری جهت خرید کالا در یک بازار رقابتی حتماً به پارامتر خدمات بعد از فروش توسط تولید کنندگان نیز توجه دارند .
۷. بررسی نظریات مصرف کنندگان ؛ مدیران بازاریابی در نهایت نقطه نظرات عکس العمل مصرف کنندگان را در مورد کالا یا خدماتی که ارائه شده می بایستی از طریق شیوه های مختلف تحقیقات بازاریابی دریافت و آنها را بررسی و مورد استفاده در تصمیم گیری و برنامه های خود قرار دهند.

## - فرآیند مدیریت بازاریابی

در مدیریت، فرآیند به معنی جریان هماهنگ و پیوسته جمع آوری اطلاعات، طبقه بندی، تجزیه و تحلیل و استنباط از آنها و بالاخره اخذ یک نتیجه منطقی از عملیات پردازش اطلاعات است.

در بازاریابی که از زیر مجموعه های مدیریت بازرگانی است، فرآیند مدیریت بازاریابی به معنی یک جریان هماهنگ و پیوسته تحقیقات بازاریابی، تعیین اهداف بلند مدت (اجرائی)، برنامه ریزی فعالیت های بازاریابی در زمینه های تولید، توزیع، پیشبرد فروش، بسته بندی، خدمات بعد از فروش و ... برای رسیدن به اهداف و بالاخره

نظارت و ارزشیابی و بحث و نقد استراتژی ها، برنامه های اجرایی و دیگر حرکات و تصمیماتی است که در طول فرآیند صورت گرفته و باعث موفقیت یا شکست برنامه های اجرایی شده است. هنگام اخذ تصمیم در زمینه بازاریابی در یک بنگاه تولیدی یا بازرگانی، علاوه بر گذراندن فرآیند تحقیقات، تعیین اهداف و برنامه ریزی باید شرایط محیطی، الزامات قانونی، وضعیت رقابت و بالاخره اوضاع و احوال مالی و اقتصادی بنگاه نیز در نظر گرفته شود. بدیهی است شرایط محیطی، قوانین حاکم، نیازهای جامعه، فشارهای رقابتی و بالاخره وضعیت سود و زیان، حجم موجودی، سرعت یا کندی بازگشت سرمایه، اقلام فروش نرفته، سفارشات موجود، تراکم یا کمبود نیروی متخصص و دیگر کاستی ها یا مزایای بنگاه در فعالیت های مدیران بازاریابی، همچون مدیران دیگر، اثر می گذارد.

## - جایگاه تحلیل های بازاریابی در تحلیل سازمانی

شناسایی نیازها، ارزش ها و عادات خرید مشتری، برای بازاریابی موفق در هر داد و ستدی امری حیاتی است. تصور کنید در صورتی که طبق یک برنامه زمانی مستمر، رفتار خرید و گرایش مشتری را بشناسید، بازاریابی شما تا چه میزان کارآمدتر می شود.

امروزه سازمان ها به منظور افزایش سهم خود در بازار رقابتی و به کارگیری بهینه منابع در ارائه خدمات متناسب با نیاز مشتریان، به بخش بندی مشتریان خود می پردازند. از این رو لازم است تا رفتار مشتریان شناسایی و تحلیل شود. مسئله اصلی این است که چگونه می توان به شناسایی و تحلیل رفتار مشتریان به منظور بخش بندی آنها و انتخاب گروه هایی از مشتریان ارزشمند برای اتخاذ استراتژی های بازاریابی هدف پرداخت.

درواقع تحقیقات بازاریابی و تحلیل مشتری در پی شناسایی و گردآوری و تجزیه و تحلیل آن دسته از اطلاعاتی برمی آید که به فرصتهای بهتری در ارائه و عرضه کالا و یا خدمات خاصی می انجامد و یا مشکلات و تهدیدهای پیش روی بازاریابی کالا و یا خدمات خاصی را تبیین کرده و راهکارهای برون رفت از مشکلات را ارائه می دهد.

## ۲-۲ Big Data و شاخصه های آن

کلان داده، به مجموعه داده هایی اطلاق می گردد که حجم و سرعت تولید آنها بیش از ظرفیت و امکانات بانک های اطلاعاتی مرسوم برای ضبط، ذخیره سازی، مدیریت و تحلیل داده است و نتوان آنها را با یک پردازشگر معمولی پردازش کرد. در واقع، عبارت Big Data مدت ها است که برای اشاره به حجم های عظیمی از داده ها که توسط سازمان های بزرگی مانند گوگل یا ناسا ذخیره و تحلیل می شوند مورد استفاده قرار می گیرد. اما به تازگی، این عبارت بیشتر برای اشاره به مجموعه های داده ای بزرگی استفاده می شود که به قدری بزرگ و حجیم هستند که با ابزارهای مدیریتی و پایگاه های داده سنتی و معمولی قابل مدیریت نیستند. مشکلات اصلی در کار با این نوع داده ها مربوط به برداشت و جمع آوری، ذخیره سازی، جست و جو، اشتراک گذاری، تحلیل و نمایش آنها است. این مبحث، به این دلیل هر روز جذابیت و مقبولیت بیشتری پیدا

می‌کند که با استفاده از تحلیل حجم‌های بیشتری از داده‌ها، می‌توان تحلیل‌های بهتر و پیشرفته‌تری را برای مقاصد مختلف، از جمله مقاصد تجاری، پزشکی و امنیتی، انجام داد و نتایج مناسب‌تری را دریافت کرد.

بیشتر تحلیل‌های مورد نیاز در پردازش داده‌های عظیم، توسط دانشمندان در علوم‌ی مانند هواشناسی، ژنتیک، شبیه‌سازی‌های پیچیده فیزیک، تحقیقات زیست‌شناسی و محیطی، جست‌وجوی اینترنت، تحلیل‌های اقتصادی و مالی و تجاری مورد استفاده قرار می‌گیرد. حجم داده‌های ذخیره‌شده در مجموعه‌های داده‌ای Big Data، عموماً به‌خاطر تولید و جمع‌آوری داده‌ها از مجموعه بزرگی از تجهیزات و ابزارهای مختلف مانند گوشی‌های موبایل، حسگرهای محیطی، لاگ نرم‌افزارهای مختلف، دوربین‌ها، میکروفون‌ها، دستگاه‌های تشخیص RFID، شبکه‌های حسگر بی‌سیم و غیره با سرعت خیره‌کننده‌ای در حال افزایش است.

امروزه به واسطه‌ی افزایش استفاده از اینترنت و شبکه‌های اجتماعی حجم داده‌های موجود در فضای اینترنت بسیار افزایش داشته است به نحوی که هرگونه پردازش به روی آن از عهده‌ی کامپیوترهای معمولی خارج است. داده‌های تولیدشده از انواع منابع با حجم بسیار زیاد سرعت بالا و ساختمان داده متفاوت را کلان داده می‌نامند. امروزه کلان داده به عنوان یک منبع غنی و ارزشمند برای رنج گسترده‌ای از سازمانها به رسمیت شناخته شده است.

کلان داده‌ها<sup>۱</sup> اصطلاحی است که به داده‌های یکپارچه با رشد سریع و ناگهانی اطلاق می‌شود و بیشتر برای توصیف مجموعه داده‌های بسیار زیاد استفاده می‌شود. در این فصل کلان داده‌ها را تعریف کرده و سیر تکامل آن را در بیست سال گذشته مورد بررسی قرار گرفته و همچنین چهار ویژگی معرف کلان داده‌ها معروف به ۴V را بیان نموده که عبارت‌اند از: حجم داده‌ها<sup>۲</sup>، تنوع داده‌ها<sup>۳</sup>، سرعت تولید داده‌ها<sup>۴</sup> و ارزش آن‌ها<sup>۵</sup>. در این فصل چالش‌های موجود در زمینه‌ی کلان داده‌ها بررسی می‌شود.

امروزه چالش اصلی همه حوزه‌های شبکه و پایگاه داده، موضوع داده‌های عظیم (کلان داده) یا Big Data است. داده‌های عظیم مفهومی است که به تازگی مطرح شده و بطور کلی به افزایش حجم اطلاعات غیرساختارمند و یکپارچه در کنار ذخیره‌سازی و پردازش آنها می‌پردازد. آنها بر روی پایگاه داده‌ها که به شکل حجیم رشد می‌کنند، ذخیره می‌شوند و ضبط، شکل‌دهی، ذخیره‌سازی، مدیریت، به اشتراک‌گذاری، تحلیل و نمایش آنها از طریق ابزارهای نوعی نرم‌افزار پایگاه داده‌ها، دشوار می‌شود.

---

<sup>۱</sup>Big data  
<sup>۲</sup>Variety  
<sup>۳</sup>velocity  
<sup>۴</sup>value

از این رو تحلیل های دقیق بر روی این داده های عظیم، منجر به تصمیم گیری های با اطمینان بیشتری شده و تصمیمات بهتر، می تواند معنای کارایی بیشتر عملیات، کاهش هزینه ها و کاهش ریسک ها باشد.

داده های بزرگ تقاضا برای متخصصان در این حوزه را به شدت بالا برده است و شرکت هایی چون IBM، Oracle، SAP، Microsoft و... بیش از ۱۵ میلیارد دلار برای توسعه نرم افزارهای مدیریت و تحلیل داده سرمایه گذاری کرده اند. داده های بزرگ نحوه کار سازمان ها و افراد را تحت تاثیر قرار می دهد. داده های بزرگ فرهنگی را در سازمان ها ایجاد می کند که از طریق آن کسب و کارها و مدیران فناوری اطلاعات را به سمت استفاده از تمامی ارزش های پنهان در داده ها سوق می دهد.

ادراک این ارزش ها به همه کارکنان سازمان ها این امکان را می دهد که با بینش وسیع تری تصمیم گیری کنند، نزدیکی بیشتری با مشتریان داشته باشند، فعالیت های خود را بهینه کنند، با تهدیدات مقابله کنند و در نهایت سرمایه های خود را بر روی منبع جدیدی از سود سرشار پنهان در داده ها متمرکز سازند. سازمان ها برای رسیدن به این مرحله نیازمند معماری جدید، ابزارهای نو و فعالیت ها و تلاش های مستمری هستند تا بتوانند از مزیت های چهارچوب های مبتنی بر داده های بزرگ بهره مند گردند.

## ۲-۲-۱ پیشینه پژوهش راجع به کلان داده ها

هفتاد سال پیش برای اولین بار تلاش برای تعیین نرخ رشد در حجم داده ها و یا آنچه مردم به عنوان "انفجار اطلاعات" روبرو شده اند، شروع شد. در این زمینه در سال ۱۹۴۴ فرمونت راید<sup>۱</sup>، کتابدار دانشگاه وسلین<sup>۲</sup>، آینده پژوهشی کتابخانه را این طور تخمین زد، که کتابخانه های دانشگاه آمریکایی هر شانزده سال، دو برابر می شوند. با توجه به این نرخ رشد، راید<sup>۳</sup> حدس زد که کتابخانه دانشگاه ییل<sup>۴</sup> در سال ۲۰۴۰ به حدود ۲۰۰،۰۰۰،۰۰۰ حجم می رسد، که بیش از ۶۰۰۰ مایل از قفسه ها را اشغال خواهد کرد، پس نیاز به یک هیات فهرست نویسی بیش از شش هزار نفر داشته و در نتیجه نیاز به منبعی که حجم بزرگی از داده ها را بتواند نگهداری کند، پیدا خواهیم کرد.

در نوامبر سال ۱۹۶۷، ب.آ.مارون<sup>۵</sup> و دکتر دمین<sup>۶</sup> فشرده سازی داده ها به صورت خودکار را در مقاله ی ارتباطات ACM، بیان کردند و "انفجار اطلاعات" که در سال های اخیر اشاره شد، ضرورتی شد تا فضای مورد نیاز ذخیره سازی برای همه ی

---

<sup>۱</sup>Fremont rider

<sup>۲</sup>Wesleyan

<sup>۳</sup>Yale

<sup>۴</sup>B. A. Maron

<sup>۵</sup>P. Demin



اطلاعات به حداقل نگه داشته شود. این مقاله بیان کرد فضای ذخیره سازی باید کاهش یافته و نرخ اطلاعات از طریق استفاده از فضای مجازی برای نگهداری اطلاعات باید افزایش یابد.

در سال ۱۹۷۵ وزارت پست و مخابرات در ژاپن، شروع به انجام اطلاعات سرشماری و ردیابی حجم اطلاعات در گردش در ژاپن کردند (این ایده برای اولین بار در یک مقاله در سال ۱۹۶۹ پیشنهاد شد). در سرشماری ۱۹۷۵ عرضه اطلاعات بسیار سریعتر از مصرف اطلاعات در حال افزایش بود.

در سال ۱۹۸۱ دفتر مرکزی آمار مجارستان، یک پروژه تحقیقاتی صنایع اطلاعات کشور را شروع کرد که حجم زیادی از اطلاعات در یک بیت قرار داشت و این بررسی ادامه داشت تا سال ۱۹۹۳، استفان دینس<sup>۱</sup> یک کتابچه راهنمای کاربر برای کمپایل یک سیستم استاندارد حساب اطلاعات ملی ایجاد کرد.

در سال ۱۹۹۰ دانشمند آمریکایی پیتر جی. دنینگ<sup>۲</sup> موضوعی به نام ذخیره تمامی بیتها را منتشر کرد. او بیان کرد که برای دانشمندان ذخیره تمام بیت ها برای حفظ وضعیت ضروری است. نرخ و حجم جریان برای جلوگیری از اختلال اطلاعات شبکه، دستگاه های ذخیره سازی، سیستم های بازیابی و همچنین به عنوان ظرفیت انسان برای درک اینکه چه ماشینی جریان داده ها را به وجود آورده و از چه ابزار نظارتی برای غربال کردن پایگاه داده باید استفاده شود زیاد است، پس باید به دنبال یک الگو و تشریح بیت هایی که قرار است ذخیره کنیم باشیم.

در ۱۹۹۷ لکس<sup>۳</sup> منتشر کرد: در جهان چه مقدار اطلاعات وجود دارد؟ او نتیجه گرفت ممکن است چند هزار پتابایت داده وجود داشته باشد و تا سال ۲۰۰۰ نوار و دیسک تولید خواهند شد تا اطلاعات را در خود نگه دارند و نظرش بر این بود که تا چند سال فقط می توانیم اطلاعات را نگهداری کنیم. بعد از آن انسان قادر به نگهداری اطلاعات نخواهد بود. بعد از آن در همان سال (۱۹۹۷) دیوید ایسورت<sup>۴</sup> در مقاله کاربرد کنترل تقاضای صفحه بندی برای خروج از هسته بیان کرد که چالش هایی برای سیستم کامپیوتری وجود دارد و آن هم وجود مجموعه داده های بسیار بزرگ است که بر روی حافظه های اصلی، دیسک محلی و دیسک های از راه دور قابل ذخیره سازی نیست. این مجموعه داده بزرگ یا Big Data نام دارد. هنگامی که مجموعه داده در حافظه اصلی جایگذاری نمی شود و یا اینکه بر روی دیسک محلی هم قابلیت ذخیره سازی نیست، راه حل رایج بدست آوردن منابع بیشتر برای ذخیره سازی می باشد. اصطلاح Big Data یا همان داده بزرگ برای اولین بار در کتابخانه دیجیتال ACM تعریف و بکار گرفته شد.

---

<sup>۱</sup>Stefan Deines

<sup>۲</sup>Peter J. Denning

<sup>۳</sup>. Lex

<sup>۴</sup>. David Eiswert.

برایسون<sup>۱</sup> و همکاران در بررسی ظاهری مجموعه گیگابایت داده در زمان واقعی در نشریه ارتباطات ACM، بیان کردند که کامپیوتر بسیار قدرتمند است و در زمینه های بسیاری از قبیل پرس و جوها، محاسبات سریع در حجم انبوهی از داده (مگابایت داده)، شبیه سازی فردی (در محدوده ۳۰۰ گیگا بایت) کاربرد دارد.

در سال ۲۰۰۰ مطالعه جامعی بر روی تعیین کمیت داده در شرایط ذخیره سازی انجام شد. داده هایی که در سالهای قبل تولید شده و به حد ۱,۵ اگزابایت رسیده بودند، در فضای بزرگی مانند رسانه های ذخیره سازی مغناطیسی ثبت و نگهداری شده اند. در کنگره جهانی اقتصادسنجی فرانسیس ایکس دایبولد<sup>۲</sup> مقاله ای تحت عنوان مدل فاکتور پویا: داده بزرگ برای اندازه گیری و پیش بینی اقتصاد کلان ارائه داده بود. او بیان کرد اطلاعات بزرگ در مقدار خود به حد انفجار رسیده اند و پیشرفت های اخیر در زمینه فن آوری ذخیره سازی و ثبت اطلاعات بسیار مؤثر بوده است.

سال ۲۰۰۵ تیم اوریلی<sup>۳</sup> در مقاله «وب ۲,۰ چیست»، بیان کرد پایگاه داده ای باید وجود داشته باشد تا تمامی اطلاعات را در خود قرار داده، سپس آنها را مدیریت کند و ما آنها را به عنوان افزار اطلاعاتی و نه به عنوان یک نرم افزار، بشناسیم.

سال ۲۰۱۰ کوکیر<sup>۴</sup> گزارشی مبتنی بر داده ها در همه جا می نویسد جهان شامل مقدار گسترده ای از اطلاعات دیجیتال است که با سرعت زیادی رو به افزایش و پیشرفت است و آنها این ویژگی داده ای را داده بزرگ نامیدند.

سال ۲۰۱۲ مجله بین المللی ارتباطات بخش ویژه با عنوان ظرفیت اطلاعات منتشر شد. آن ها در این مقاله بیان داشتند عرضه رسانه ها در اماکن مسکونی خصوصاً در آمریکا از سال ۱۹۶۰ تا ۲۰۰۵ افزایش یافته است و اگر نسبت عرضه به تقاضا به همین صورت پیش برود به حدود سه برابر داده ای که در سال ۲۰۰۳ در اختیار کاربران قرار داشت، می رسیم.

گفتار دوم: سیر تکامل کلان داده ها

در بیست سال گذشته در حوزه های مختلف، حجم داده ها به مقدار زیادی افزایش یافته است. طبق گزارش موسسه ی بین المللی داده (IDC<sup>۵</sup>)، در سال ۲۰۱۱ حجم داده های ایجاد و تکثیر شده در سراسر دنیا ۱/۸ زتابایت (۱۰ به توان ۲۱ بایت) بود که این مقدار در ظرف پنج سال نزدیک به ۹ برابر شده است (گانتز و رینسل، ۲۰۱۱).

اصطلاح کلان داده ها به داده های یکپارچه با رشد سریع و ناگهانی اطلاق شده و بیشتر برای توصیف مجموعه داده ها ی بسیار زیاد استفاده میشود. در مقایسه با مجموعه داده های سنتی، کلان داده ها معمولاً شامل انبوهی از داده های غیر ساخت یافته است که اغلب نیاز به تجزیه و تحلیل بی درنگ دارند. کلان داده ها همچنین فرصت های نوینی برای کشف

<sup>۱</sup>Bryson

<sup>۲</sup>Francis x. Diebold

<sup>۳</sup>O'Reilly

<sup>۴</sup>Cookier

<sup>۵</sup>.International Data Corporation

مقادیر جدید فراهم نموده و ما را برای فهم عمیق ارزش های پنهان یاری می کنند اما چالش هایی نظیر چگونگی مدیریت و سازماندهی چنین داده هایی را به وجود می آورد. در حال حاضر، کلان داده ها توجه بسیاری از صنایع ، دانشگاه ها و سازمان های دولتی را به خود جلب کرده است. به عنوان مثال، موضوعات مربوط به کلان داده ها معمولاً در رسانه های عمومی مثل *Nature*، *Science*، *New York Times*، *Economist* و *National Public Radio* پوشش داده میشود (کوکی، ۲۰۱۰). دو مجله ی برتر علمی *Nature* و *Science* نیز در مورد اهمیت و چالش های کلان داده ها ستون های ویژه ای را به آن اختصاص داده اند. سازمان های دولتی بسیاری برای شتاب دادن به تحقیقات و کاربرد های کلان داده ها از طرح های بزرگی خبر داده و صنایع نیز اخیراً مجذوب پتانسیل بالای کلان داده ها شده اند. بدون تردید عصر کلان داده ها در حال پیدایش است (لوهر، ۲۰۱۲).

رشد سریع اخیر کلان داده ها ، به طور عمده ناشی از زندگی روزانه ی افراد به ویژه در ارتباط با خدمات شرکت های اینترنتی است. به عنوان مثال در هر صد ها پتابایت پردازش داده در گوگل و بالغ بر ده پتابایت داده ثبت وقایع در فیس بوک تولید میشود . بایدو<sup>۱</sup> یک شرکت چینی ، ده ها پتابایت و یکی از شرکت های تابعه ی علی بابا<sup>۲</sup> ده ها ترابایت داده را روزانه در کسب و کارهای اینترنتی خود تولید و پردازش می کنند. به دلیل رشد سریع مجموعه ی کلان داده ها، مشکلات چالش برانگیزی به وجود آمده است که به حل فوری آن ها نیاز مبرم وجود دارد . اول ، پیشرفت های اخیر فناوری اطلاعات تولید داده را آسان تر کرده است ، به طور متوسط در هر دقیقه هفتاد و دو ساعت فیلم در یوتیوب<sup>۳</sup> بارگذاری شده است (مایر، ۲۰۱۳) . بنابراین اول با چالش بزرگی در جمع آوری و یکپارچه کردن انبوهی از داده ها از منابع زیاد و توزیع شده مواجه هستیم . دوم ، داده های جمع شده که به طور فزاینده ای در حال رشد است منجر به مشکل جدی در چگونگی ذخیره و مدیریت مجموعه داده های ناهمگن و عظیمی با نیازمندی های خاص روی زیر ساخت های نرم افزاری و سخت افزاری می شود . سوم ، در مقابل ناهمگنی ، مقیاس پذیری ، بی درنگ بودن ، پیچیدگی و محرمانگی کلان داده ها ، ما باید مجموعه داده ها را به گونه ای موثر با استفاده از روش های تحلیل ، مدل سازی ، مصور سازی<sup>۴</sup>، پیش بینی و بهینه سازی از سطوح مختلف استخراج کنیم تا به خصوصیات ذاتی آن ها دست پیدا کنیم و پروسه ی تصمیم گیری را بهبود بخشیم.

رشد سریع رایانش ابری<sup>۵</sup> و اینترنت اشیا<sup>۶</sup> (IoT) ، رشد داده ها را شدت زیادی بخشیده است. رایانش ابری، کانال ها و سایت های قابل دسترس امنی را برای داده های پر ارزش فراهم می کند. اینترنت اشیا با حسگر هایی در سراسر دنیا، داده های ذخیره و پردازش شده در ابر را جمع آوری کرده و انتقال می دهد. این داده ها هم از لحاظ کمیت و هم روابط متقابل<sup>۷</sup>

---

<sup>۱</sup>Baidu

<sup>۲</sup>AliBaba

<sup>۳</sup>You Tube

<sup>۴</sup>Visualization

<sup>۵</sup>Cloud computing

<sup>۶</sup>Internet of things

<sup>۷</sup>Mutual relations

از ظرفیت زیر ساخت ها و معماری های فناوری اطلاعات در موسسات موجود افزون تر است و نیاز به بی درنگ بودن داده ها، به ظرفیت محاسباتی موجود فشار زیادی وارد خواهد کرد.

## ۲-۲ معرفی کلان داده ها و خصوصیات آن ها

کلان داده ها یک مفهوم انتزاعی است. صرف نظر از حجم داده ها ، کلان داده ها خصوصیات دیگری دارد که تفاوت آن را با داده های انبوه<sup>۱</sup> و داده های خیلی بزرگ<sup>۲</sup> مشخص می کند. اگر چه امروزه اهمیت کلان داده ها معمولاً شناخته شده است اما هنوز برای تعریف آن نظرات متفاوتی وجود دارد. معمولاً، کلان داده ها به مجموعه داده هایی گفته می شود که توانایی دریافت<sup>۳</sup>، اکتساب<sup>۴</sup>، مدیریت و پردازش آن ها در یک زمان قابل قبول به وسیله فناوری اطلاعات و ابزارهای نرم افزاری/ سخت افزاری سنتی وجود ندارد. بدین دلیل شرکت ها ، موسسات علمی و صنعتی، پژوهشگران ممتاز ، تحلیل گران داده و کارورزان حرفه ای مختلف تعریف های متفاوتی برای کلان داده ها دارند. تعاریفی که در ادامه می آید برای فهم بهتر معنی اجتماعی، اقتصادی و فنی کلان داده ها به ما کمک می کند.

در سال ۲۰۱۰، هادوپ آپاچی<sup>۵</sup> کلان داده ها را این گونه تعریف کرد: " مجموعه داده هایی که اکتساب، مدیریت و پردازش آن ها به وسیله کامپیوتر های معمولی در یک بازه ی زمانی قابل قبول امکان پذیر نیست." بر اساس این تعریف، در ماه مه ۲۰۱۱، مک کینزی و کمپانی<sup>۶</sup> گه یک آژانس مشاور جهانی است کلان داده را این گونه معرفی کرد: " مرز بعدی برای نوآوری، رقابت و بهره وری . "کلان داده ها باید مجموعه داده هایی را پوشش دهد که امکان اکتساب، ذخیره و مدیریت آن ها به وسیله ی نرم افزار های پایگاه داده ی کلاسیک وجود ندارد. این تعریف دو معنی ضمنی را در بر می گیرد : اول، حجم مجموعه داده ها با استاندارد در حال تغییر کلان داده ها مطابق است و ممکن است با زمان یا با پیشرفت های فناوری رشد کند. دوم، حجم مجموعه داده ها با استاندارد کلان داده ها در کاربردهای گوناگونی که هر یک با دیگری متفاوت است باید مطابق باشد . در حال حاضر محدوده ی کلان داده ها معمولاً از چندین ترابایت تا چندین پتابایت است . با این تعریف مک کینزی و کمپانی، به نظر می رسد حجم داده ها به تنهایی معیاری برای کلان داده نیست. دو ویژگی کلیدی دیگر آن، رشد فرایند میزان داده ها و مدیریتی است که به وسیله ی فناوری های پایگاه داده ی سنتی قابل انجام نیست.

---

<sup>۱</sup>Massive data

<sup>۲</sup>Very big data

<sup>۳</sup>Perceived

<sup>۴</sup>Acquired

<sup>۵</sup>Apache Hadoop

<sup>۶</sup>Mckinsey & Company

حقیقت امر این است که کلان داده ها از سال ۲۰۰۱ تعریف شده بود. داوگ لانی<sup>۲</sup> تحلیل گری از متا<sup>۳</sup> (گارتنر فعلی) در یک گزارش تحقیقاتی، فرصت ها و چالش های ایجاد شده از داده های افزایش یافته را با مدل ۳<sup>۴</sup> (یعنی افزایش حجم، سرعت و تنوع) تعریف کرد (لانی، ۲۰۰۱). گرچه در ابتدا این مدل برای تعریف کلان داده ها استفاده نشد، اما گارتنر و شرکت های دیگری مثل آی بی ام و بعضی سازمان های تحقیقاتی میکروسافت طی ۱۰ سال از این مدل برای تعریف کلان داده ها استفاده کرده اند (بیر، ۲۰۱۱). در مدل ۳<sup>۴</sup>، حجم یعنی با تولید و جمع آوری داده های انبوه، اندازه داده ها به طور فزاینده ای حجیم میشود. سرعت به معنی به موقع بودن کلان داده ها به ویژه در جمع آوری و تحلیل و غیره است، داده ها باید جریان سریع و به موقعی داشته باشد. در واقع سرعت به معنی بهره وری بیشینه از ارزش تجاری کلان داده ها است. تنوع، انواع گوناگون داده ها را نشان میدهد که داده هایی به شکل نیمه ساخت یافته<sup>۵</sup> و غیر ساخت یافته<sup>۶</sup> مثل صدا، فیلم، صفحات وب، متن و همچنین داده های ساخت یافته<sup>۷</sup> سنتی را دربر می گیرد (زیکوپولوس و همکاران، ۲۰۱۱).

IDC یکی از تاثیرگذارترین رهبران در کلان داده ها و زمینه های تحقیقاتی آن نظر متفاوتی دارد. در سال ۲۰۱۱ آن ها در یک گزارش تحقیقاتی کلان داده ها را این گونه تعریف کردند: " فناوری های کلان داده ها نسل جدیدی از فناوری ها و معماری را توصیف می کند که برای استخراج ارزش اقتصادی از حجم بسیار بزرگی از انواع گوناگون داده با سرعت بالا در اکتساب، کشف یا تحلیل داده طراحی شده است." در این تعریف ویژگی های کلان داده ها با چهار V خلاصه میشود که یعنی حجم (حجم عظیم)، تنوع (انواع گوناگون)، سرعت (سرعت تولید) و ارزش (ارزش کلان اما با چگالی کم). این تعریف ۴<sup>۸</sup> مورد استقبال قرار گرفت زیرا معنی و ضرورت کلان داده ها ارزش های پنهان زیادی را مشخص می کرد. همچنین این تعریف به مسائل مهم و حیاتی کلان داده ها یعنی چگونگی کشف ارزش ها از مجموعه داده هایی با مقیاس بزرگ، انواع گوناگون و تولید سریع آن ها اشاره می کند (میژر، ۲۰۱۱).

جی پاریک<sup>۹</sup> معاون سر مهندسان فیس بوک گفت: " اگر شما از داده های جمع آوری شده استفاده نکنید فقط می توانید مالک بخشی از داده ها به غیر از کلان داده ها باشید " (مایر، ۲۰۱۳). موسسه ی ملی استاندارد ها و فناوری ایالات متحده<sup>۱۰</sup> کلان داده ها را این گونه تعریف می کند: " کلان داد ها بیانگر داده هایی هستند که حجم داده ها، سرعت اکتساب داده ها یا نمایش داده ها، ظرفیت استفاده از روش های رابطه ای سنتی را برای تحلیل موثر محدود می کند یا داده

<sup>۲</sup>Doug Laney

<sup>۳</sup>META

<sup>۴</sup>Semi-structured

<sup>۵</sup>Unstructured

<sup>۶</sup>Structured

<sup>۷</sup>Great volume

<sup>۸</sup>Various modalities

<sup>۹</sup>Rapid generation

<sup>۱۰</sup>Huge value but very low density

<sup>۱۱</sup>Jay Parikh

<sup>۱۲</sup>NIST

هایی که با فناوری های درشت نمایی افقی<sup>۱</sup> مهم ممکن است به صورت کارا پردازش شده باشند"، این تعریف روی جنبه های فناوری کلان داده ها متمرکز است و برای تحلیل و پردازش کلان داده ها ضرورت توسعه و استفاده از فناوری ها و روش های کارا را نشان می دهد.

بین صنعت و دانشگاه بحث های زیادی در مورد تعریف کلان داده ها وجود داشته است (تیم Or، ۲۰۱۱).

## ۲-۲-۲ ارزش کلان داده ها :

مک کینزی و کمپانی بعد از تحقیقات عمیقی بر روی پنج هسته ی (الف) سیستم سلامت ایالات متحده (ب) مدیریت بخش عمومی اروپا (پ) خرده فروشی های ایالات متحده ، (ت) کارخانه داران جهانی و (ث) داده های مکانی اشخاص به صورت جهانی نظر خود را درباره چگونگی ارزش ایجاد شده توسط کلان داده ها بیان کردند. با تحقیق روی این پنج هسته نمایش دهنده ی اقتصاد جهانی، گزارش مک کینزی نشان داد که کلان داده ها ممکن است نقش اساسی در عملکرد اقتصادی بازی کنند، برای بخش های عمومی و موسسات بهروری و رقابتی بودن را بهبود بخشند و برای مصرف کننده فواید زیادی ایجاد کنند. مک کینزی در ارزش ایجاد شده توسط کلان داده ها را چنین خلاصه کرد: اگر کلان داده ها می توانست به صورتی خلاقانه و موثر برای بهبود کیفیت و کارایی استفاده شود ارزش بالقوه حاصل از داده ها در صنعت دارویی ایالات متحده ممکن بود از ۳۰۰ میلیارد دلار آمریکا عبور کند، هزینه ی صرف شده ی سیستم سلامت ایالات متحده را بیشتر از ۸ درصد کاهش دهد و سود خرده فروشانی که کاملاً از کلان داده ها استفاده میکنند به بیشتر از ۶۰ درصد خواهد رسید. کلان داده ها همچنین ممکن است برای بهبود کارایی کاربرد های دولتی استفاده شود مثل اقتصاد توسعه یافته ی اروپا که با پیشگیری از تاثیر کلاه برداران، خطاها و اختلاف مالیاتی می تواند ۱۰۰ میلیارد یورو پس انداز کند (مانیکا، ۲۰۱۱).

گزارش مک کینزی به عنوان یک پیش بینی مربوط به آینده مورد توجه قرار گرفت. حقایق پیش رو ممکن است ارزش کلان داده ها را تصدیق کنند. در آنفولانزای همه گیر ۲۰۰۹، گوگل اطلاعات به موقعی را با تحلیل کلان داده ها فراهم کرد که ارزش اطلاعات گوگل حتی از اطلاعات فراهم شده توسط مراکز پیشگیری بیماری بیشتر بود. تقریباً در همه کشورها، بیمارستان ها باید نوع جدید آنفولانزا را به موسساتی مثل مراکز پیشگیری بیماری گزارش میکردند. اما بیماران معمولاً به محض سرایت بیماری به پزشک مراجعه نمی کردند. به علاوه، فرستادن اطلاعات از بیمارستان ها به مراکز پیشگیری بیماری همچنین تحلیل این اطلاعات توسط این مراکز طول می کشید. بنابراین هنگامی مردم از یک آنفولانزای همه گیر مطلع می شدند که ممکن بود از شیوع بیماری با نوعی بیهوشی خطرناک، یک یا دو هفته ی گذشته باشد. گوگل دریافت که در موتورهای جستجو تعداد مدخل جستجوی اطلاعات در هنگام شیوع آنفولانزا نسبت به زمان های معمولی متفاوت و میزان استفاده از این اطلاعات وابسته به زمان و مکان شیوع آنفولانزا بود. گوگل چهل و پنج دسته مدخل

---

<sup>۱</sup>Horizontal zoom technologies

جستجوی اطلاعات مرتبط با شیوع آنفولانزا پیدا کرده و آن‌ها را در مدل‌های ریاضی خاصی ترکیب نمود تا شیوع آنفولانزا را پیش‌بینی و حتی مکان‌هایی را پیش‌بینی کند که آنفولانزا آن‌جا شیوع پیدا خواهد کرد. نتایج این تحقیق در *Nature* منتشر شد (گینسبرگ، ۲۰۱۱).

در سال ۲۰۰۸، یک شرکت اقتصادی سای – تک<sup>۱</sup> در ایالت متحده ی آمریکا (پیش‌بینی خرید مایکروسافت) از داشتن یک سیستم پیش‌بینی بلیط هواپیمای مسافربری خبر داد که روند و میزان افزایش یا کاهش قیمت بلیط را پیش‌بینی می‌کند. این سیستم با موتور جستجوی بینگ<sup>۲</sup> مایکروسافت به هم پیوسته بودند. در سال ۲۰۱۲ این سیستم در هزینه ی بلیط هر مسافر نزدیک به پنجاه دلار صرفه جویی کرد که دقت پیش‌بینی آن بیشتر از ۷۵ درصد بود.

امروزه، داده‌ها یک عامل بهره‌وری مهم بوده که با مواد اولیه و سرمایه‌ی انسانی قابل مقایسه است. همان‌طور که چند رسانه‌ای، رسانه‌ی اجتماعی و اینترنت اشیا به سرعت در حال رشد هستند شرکت‌ها اطلاعات زیادی را جمع می‌کنند که منجر به رشد نمایی حجم داده‌ها می‌شود. کلان‌داده‌ها در ایجاد ارزش برای تجارت و مصرف‌کنندگان، پتانسیل بزرگ و فزاینده‌ای خواهد داشت.

بازاریابی مدرن وابستگی بسیار زیادی به اطلاعات دارد. در این میان اهمیت کلان‌داده‌ها (Big Data) هر روز بیشتر از گذشته می‌شود.

کلان‌داده (Big Data) وعده‌های زیادی به صنعت بازاریابی داده است. اصلی‌ترین وعده، پاسخ به دو سوال اساسی است. اول اینکه با کمک آن‌ها بتوان فهمید چه کسی، چه محصولی را در چه زمانی با چه قیمتی می‌خرد؟ و دوم اینکه چگونه می‌توان بین محصولی که مشتریان می‌بینند یا در موردش شنیده یا می‌خوانند، و چیزی که نهایتاً می‌خرند یا مصرف می‌کنند، ارتباطی منطقی پیدا کرد؟

پاسخ به این دو سوال، بازاریابی را بسیار اثرگذارتر از روش‌هایی مثل هدفگیری و شناسایی مشتری می‌کند؛ روش‌هایی که نیمی از بودجه بازاریابی را هدر می‌دهند.

بازاریاب‌ها برای یافتن پاسخ باید آموزش ببینند تا بتوانند از کلان‌داده‌ها در این مسیر و برای پیش‌بینی خرید بعدی مشتریان، بهره‌برند. در این راه، بازاریاب‌ها باید تلاش کنند در تشخیص الگوی مصرف هر مشتری و شناخت علاقمندی‌ها، خواسته‌ها و عادات خرید و مصرف او، موفق و بهینه عمل کنند. نتیجه، جزییاتی دقیق و کامل از عملکرد مصرف‌کننده‌ها و امکان پیش‌بینی خریدهای بعدی آن‌ها خواهد بود.

---

<sup>۱</sup>Sci-tech

<sup>۲</sup>Bing

اما بسیاری از صنایع در تلاش برای تشخیص سریع‌تر اهداف مشتری‌ها، با حقیقتی نگران‌کننده روبرو شده‌اند. تکنیک‌های مورد نظر برای پیش‌بینی خریدها و جذب مشتریان، تنها برای بازه‌های کوتاه‌مدت نتیجه‌بخش است. در این بازی، رقابت بازاریاب‌ها برای شناسایی حرکت بعدی مشتریان به توازن رسیده و تلاش‌ها بر کسب سود حاشیه‌ای متمرکز شده‌اند. نتیجه، رقابتی بی‌برنده در بازه‌ی زمانی کوتاه مدت است. همه در بازه‌های زمانی میان مدت و بلندمدت دوشادوش هم در حرکت به سوی برابری پیش رفته و هیچ‌یک از مدعیان مزیت رقابتی پایداری برای پیروزی در اختیار ندارند.

البته معنی گفته‌های بالا، توقف تلاش شرکت‌ها برای پیش‌بینی بهینه‌تر رفتار مشتریان نیست؛ اما زمانی که همه‌ی رقبا مشغول همین تلاش هستند و در شرایطی که ماندن در بازی هم پاداش‌های خودش را برای صنایع مختلف دارد، حداکثر نتیجه‌ای که از این تلاش به دست می‌آید، کسب درآمد بالاتر از حد میانگین خواهد بود. در بسیاری صنایع از جمله گردشگری، بیمه، ارتباطات، موسیقی و حتی خودروسازی، تکنیک‌های بازاریابی به حدی پیشرفت کرده که رقبا به سطحی برابر در قدرت جذب مشتری رسیده‌اند. مهم‌ترین و تنها مولفه‌ی باقیمانده برای جذب مشتری بیشتر، قدرت پیش‌بینی رفتار مصرف‌کننده‌های نهایی است.

برای دستیابی به این مزیت رقابتی، برنامه‌های بازاریابی، به تحلیل اطلاعات کلان داده‌ها برای پاسخ به سوالات استراتژیک درباره‌ی چگونگی جذب بیشتر مشتریان و ایجاد حس وفاداری و برقراری ارتباط پایدار با آن‌ها، نیاز دارند؛ ارتباطی که بتواند با پایداری در بلندمدت، نیازهای مشتریان را در مقابل رقبا، مرتفع کند.

سوال‌های استراتژیک یادشده، به موارد مهمی همچون کشف روش‌های ایجاد حس وفاداری در مشتریان کمک شایانی می‌کنند. در شرایط فعلی، قیمت‌گذاری مناسب برای جذب بیشتر مشتری، کارساز نیست، بلکه باید راهی برای ایجاد ارزش بیشتر در محصولات و خدمات پیدا کرد که مشتریان قدر آن را بدانند. یافتن روش‌هایی برای عدم رغبت مشتریان به امتحان کردن یا انتخاب محصول و خدمات رقیب، استراتژی‌ای بسیار مهم‌تر از تبدیل مشتریان رقیب به مصرف‌کنندگان کالاها یا خدمات خود است.

پاسخ به این سوالات اساسی و یافتن راه‌حل‌های مطلوب، نیازمند استفاده‌ی متفاوت از کلان داده‌ها است. لازم است به جای یافتن راه‌حل‌هایی برای هدف‌گیری مشتریان بالقوه، به دنبال استفاده از کلان داده‌ها جهت کسب صفات ارزشمند در محصولات و خدمات بود. این صفات ارزشمند، چیزهایی هستند که مشتریان برای انتخاب گزینه‌ی ایده‌آل به آن توجه می‌کنند. سوال اساسی و متفاوت این است: کلان داده‌ها چگونه می‌توانند به ما و نه مشتریان کمک کنند؟

با کمک کلان داده‌ها می‌توان اطلاعاتی جدید را برای ایجاد محصولات و خدماتی کاملاً تازه کشف کرد. مثال ساده در همین رابطه، استفاده از کلان داده‌ها برای کمک به تسریع روند یافتن نتیجه‌ی مورد نظر کاربران به هنگام استفاده از موتورهای جست‌وجو می‌باشد. روشی که در حال حاضر شرکت‌هایی مانند آمازون و نتفلیکس برای کمک به کاربران‌شان



جهت یافتن سریع‌تر گزینه‌های مورد نظرشان، از آن بهره می‌برند. در مثالی دیگر می‌توان به تجمیع منابع اطلاعاتی که شرکت‌ها از نتیجه‌های جست‌وجوی سایر مشتریان و کاربران‌شان به‌دست می‌آورند اشاره کرد. اطلاعاتی که با کمک آن‌ها می‌توان با تشخیص روند جست‌وجوی سایر مشتریان و مقایسه‌ی آن با مشتری فعلی، نتایج بهتری را در اختیار او گذاشت.

نگاهی به استارت‌آپ‌هایی که امکان استخراج اطلاعات قابل انکاتری از کلان داده‌ها فراهم کرده‌اند، می‌تواند جذاب باشد.

شرکت Opower به کاربران اجازه می‌دهد با به اشتراک‌گذاری صورت حساب‌هایشان با دوستان فیسبوکی خود، رتبه‌ی خود را در قیاس با دیگران مشاهده کنند. INRIX با تجمیع میزان ترافیک مصرفی کاربران، امکان گزارش‌گیری بی‌درنگ را مهیا کرده است. شرکت شناخته شده‌ی Zillow با ادغام اطلاعات از منابع مختلف، آماری تلفیقی از ویژگی‌ها، قیمت و سایر صفات موثر در معاملات املاک را در اختیار خریداران، فروشندگان و معامله‌گرهای ملک قرار می‌دهد.

این شرکت‌ها، استفاده از کلان داده‌ها را سنگ‌بنای کسب و کارشان قرار داده‌اند و مسیر موفقیت‌شان باید الگویی برای تمامی شرکت‌ها در حوزه‌های مختلف باشد. چرا که امروزه هیچ کسب و کاری نمی‌تواند بدون استفاده‌ی موثر از داده‌ها، به حیات خود ادامه دهد.

هر شرکت باید سه سوال مهم زیر را برای ماندن در بازار رقابت در نظر بگیرد:

۱. چه نوع اطلاعاتی می‌تواند هزینه و ریسک مشتریان نهایی ما را به حداقل برساند؟

شرکت‌های بزرگی مانند Amazon، Netflix، eBay، Uber، TripAdvisor، Zagat، Yelp اهمیت زیادی به جمع‌آوری اطلاعات مرتبط با رتبه‌بندی تامین‌کننده‌های کالا و خدمات می‌دهند. چرا که این اطلاعات میزان ریسک‌پذیری مشتریان‌شان را به حداقل می‌رساند. در حال حاضر این دست اطلاعات، فاکتور مهمی برای ارزیابی عملکرد تامین‌کننده‌ها به حساب می‌آیند. با این حال، مشتریان به طرز فزاینده‌ای به اطلاعاتی که کمک کند بتوانند رفتار کاربرانی مثل خودشان را در مقابله با محصولات و خدمات مختلف ببینند، توجه نشان می‌دهند. در اختیار گذاشتن چنین اطلاعاتی به قدرت درک عمیق رفتار مشتریان و کشف الگوی توجه‌شان به عملکرد یکدیگر نیاز دارد؛ و این چیزی است که با کمک کلان داده‌ها می‌توان به آن دست یافت.

۲. چه نوع اطلاعاتی در دسترس هستند؟ آیا باید آن‌ها را جمع‌آوری و مورد استفاده قرار داد؟

آیا جمع‌آوری اطلاعاتی حاشیه‌ای مانند موقعیت مکانی کاربران می‌تواند ارزشمند باشد؟ این چیزی است که InVenture که یک استارت‌آپ جذاب در آفریقا است، به آن پرداخته است. این شرکت با تبدیل اطلاعاتی که به‌صورت اتفاقی در تلفن همراه کاربران تولید می‌شود، امکان تشخیص مشتریانی را که به خدمات مالی خاصی دسترسی یا نیاز داشته یا دارند برقرار کرده است. در محیطی که هیچ تاریخچه‌ای از وضعیت اعتبار مالی بخش اعظمی از جمعیت وجود ندارد، استفاده از اطلاعات

حاشیه‌ای که به‌ظاهر بی‌اهمیت به نظر می‌رسند، به دستیاری مناسب برای موسسات مالی در جهت تصمیم‌گیری برای پرداخت یا عدم پرداخت وام به مشتریان تبدیل شده است. برای مثال این شرکت تشخیص داده که احتمال بازپرداخت منظم و کامل وام افرادی که در دفترچه تماس تلفن همراه‌شان نام و نام‌خانوادگی افراد را به طور کامل ثبت کرده‌اند، بیشتر است.

۳. تنوع و اختلاف سلیقه‌ی مشتریان، چگونه بر میزان سود حاصل از جمع‌بندی اطلاعات سایرین تاثیر می‌گذارد؟

برای مثال شرکتی که محصولات مرتبط با کشاورزی مانند بذر، کود و آفت‌کش را عرضه می‌کند، می‌تواند با جمع‌آوری اطلاعات خرید کشاورزان تشخیص دهد که هر کشاورز بر اساس نوع زمین و شرایطش، چه خرید بهینه‌ای را می‌تواند انجام دهد. جمع‌آوری اطلاعات تعداد زیادی کشاورز در کنار طبقه‌بندی مولفه‌هایی مانند نوع خاک، آب‌وهوا و شرایط زمین هر یک از آن‌ها، می‌تواند کمک کند که برای هر کشاورز، بسته‌ی بهینه‌ای از محصولات پیشنهاد شود و در نهایت نتیجه‌ی بهتری نیز برداشت شود.

کلان داده‌ها می‌توانند بسیاری از سوالات اساسی بازاریاب‌ها را پاسخ دهند. یافتن پاسخ سوالاتی که می‌تواند برای مدت‌ها برای آن‌ها بی‌جواب مانده باشد با کمک گرفتن و تحلیل صحیح کلان داده‌ها امکان‌پذیر می‌شود. این، تنها راه‌حل برای صنعت بازاریابی است که با کمک مجموعه‌ای از اطلاعات، بهترین بازخورد و بالاترین سطح بازدهی را در درازمدت برای صنایع مختلف، حاصل کند.

## ۲-۲-۴ چالش‌های کلان داده‌ها

در بحث کلان داده، ما نیاز داریم که داده‌ها را به منظور استخراج اطلاعات، کشف دانش و در نهایت تصمیم‌گیری در خصوص مسائل مختلف کاربردی به صورت صحیح مدیریت کنیم. مدیریت داده‌ها عموماً شامل ۵ فعالیت اصلی می‌باشد:

۱. جمع‌آوری
۲. ذخیره‌سازی
۳. جستجو
۴. به اشتراک گذاری
۵. تحلیل

تا کنون چالشهای زیادی در حوزه کلان داده مطرح شده است که تا حدودی از جنبه تئوری ابعاد مختلفی از مشکلات این حوزه را بیان میکنند. این چالش ها در ابتدا سه بعد اصلی حجم داده، نرخ تولید و تنوع به عنوان  $V'S^3$  مطرح شدند ولی در ادامه چالش های بیشتری در ادبیات موضوع توسط محققان مطرح شده است [۱،۲] خواهیم پرداخت:

- حجم داده (Volume): حجم داده ها به صورت نمایی در حال رشد می باشد. منابع مختلفی نظیر شبکه های اجتماعی، لاگ سرورهای وب، جریان های ترافیک، تصاویر ماهواره ای، جریان های صوتی، تراکنش های بانکی، محتوای صفحات وب، اسناد دولتی و ... وجود دارد که حجم داده بسیار زیادی تولید می کنند.
- نرخ تولید : داده ها از طریق برنامه های کاربردی و سنسورهای بسیار زیادی که در محیط وجود دارند با سرعت بسیار زیاد و به صورت بلادرنگ تولید می شوند. بسیاری از کاربردها نیاز دارند به محض ورود داده به درخواست کاربر پاسخ دهند. ممکن است در برخی موارد نتوانیم به اندازه کافی صبر کنیم تا مثلاً یک گزارش در سیستم برای مدت طولانی پردازش شود.
- تنوع : انواع منابع داده و تنوع در نوع داده بسیار زیاد می باشد که در نتیجه ساختارهای داده ای بسیار زیادی وجود دارد. مثلاً در وب، افراد از نرم افزارها و مرورگرهای مختلفی برای ارسال اطلاعات استفاده می کنند. بسیاری از اطلاعات مستقیماً از انسان دریافت میشود و بنابراین وجود خطا اجتناب ناپذیر است. این تنوع سبب میشود جامعیت داده تحت تاثیر قرار بگیرد. زیرا هرچه تنوع بیشتری وجود داشته باشد، احتمال بروز خطای بیشتری نیز وجود خواهد داشت.
- صحت : با توجه به اینکه داده ها از منابع مختلف دریافت میشوند، ممکن است نتوان به همه آنها اعتماد کرد. مثلاً در یک شبکه اجتماعی، ممکن است نظرهای زیادی در خصوص یک موضوع خاص ارائه شود. اما اینکه آیا همه آنها صحیح و قابل اطمینان هستند، موضوعی است که نمیتوان به سادگی از کنار آن در حجم بسیار زیادی از اطلاعات گذشت. البته بعضی از تحقیقات این چالش را به معنای حفظ همه مشخصه های داده اصلی بیان کرده اند که باید حفظ شود تا بتوان کیفیت و صحت داده را تضمین کرد. البته تعریف دوم در مولدهای کلان داده صدق میکند تا بتوان داده ای تولید کرد که نشان دهنده ویژگی های داده اصلی باشد.
- اعتبار : با فرض اینکه دیتا صحیح باشد، ممکن است برای برخی کاربردها مناسب نباشد یا به عبارت دیگر از اعتبار کافی برای استفاده در برخی از کاربردها برخوردار نباشد.
- نوسان : سرعت تغییر ارزش داده های مختلف در طول زمان میتواند متفاوت باشد. در یک سیستم معمولی تجارت الکترونیک، سرعت نوسان داده ها زیاد نیست و ممکن است داده های موجود مثلاً برای یک سال ارزش خود را حفظ کنند،

اما در کاربردهایی نظیر تحلیل ارز و بورس، داده با نوسان زیادی مواجه هستند و داده ها به سرعت ارزش خود را از دست میدهند و مقادیر جدیدی به خود می گیرند. اگرچه نگهداری اطلاعات در زمان طولانی به منظور تحلیل تغییرات و نوسان داده ها حائز اهمیت است. افزایش دوره نگهداری اطلاعات، مسلماً هزینه های پیاده سازی زیادی را دربر خواهد داشت که باید در نظر گرفته شود.

- نمایش : یکی از کارهای مشکل در حوزه کلان داده، نمایش اطلاعات است. اینکه بخواهیم کاری کنیم که حجم عظیم اطلاعات با ارتباطات پیچیده، به خوبی قابل فهم و قابل مطالعه باشد از طریق روش های تحلیلی و بصری سازی مناسب اطلاعات امکان پذیری است.

- ارزش : این موضوع دلالت بر این دارد که از نظر اطلاعاتی برای تصمیم گیری چقدر داده حائز ارزش است. به عبارت دیگر آیا هزینه ای که برای نگهداری داده و پردازش آنها میشود، ارزش آن را از نظر تصمیم گیری دارد یا نه. معمولاً داده ها میتوانند در لایه های مختلف جابجا شوند. لایه های بالاتر به معنای ارزش بیشتر داده می باشند. بنابراین برخی از سازمانها میتوانند هزینه بالای نگهداری مربوط به لایه های بالاتر را قبول کنند.

گزارش موسسه تحقیقاتی مکنزی :موسسه مکنزی در گزارشی که در سال ۲۰۱۱ و با عنوان "مرز بعدی برای نوآوری، رقابت و بهره وری" منتشر کرده است به برخی از واقعیت های بسیار مهم در حوزه داده های عظیم اشاره کرده است که با توجه به اهمیت آن به برخی از نکات مهم اشاره می گردد .

- فراگیر شدن داده در هر صنعت و کسب و کار بگونه ای که هم اینک داده به یکی از اجزای مهم و کلیدی در چرخه تولید و عرضه محصول و یا سرویس تبدیل شده است: گروه های مطالعاتی متعددی بر روی این موضوع کار کرده اند که چه میزان داده در سطح جهان تولید ، ذخیره و مصرف می شود. با این که حوزه کاری آنها با یکدیگر متفاوت بوده است و هر یک به نتایج خاصی در آن حوزه بخصوص دست یافته اند ولی تمامی آنها به رشد نمایی داده درسالیان اخیر بطور مشترک اشاره داشته اند . رشد نمایی داده همچنان به روند خود ادامه خواهد داد و با بکارگیری فناوری ها ، پلت فرم ها و قابلیت های تحلیل جهت برخورد با داده ، شاهد نوآوری و خلاقیت های متعددی در این حوزه خواهیم بود . تحلیل رفتار مشتریان در زندگی دیجیتالی هر یک ، نمونه ای از دستاوردهای مهم درحوزه مدیریت و تحلیل صحیح داده های عظیم در هر بنگاه اقتصادی است .

- داده های عظیم قادر به تولید ارزش با روش های مختلفی می باشند : در این گزارش به پنج روش اشاره شده است که سازمان ها می توانند با بکارگیری هر یک از آنها در مسیر خلق ارزش های جدید برای کسب و کار خود اقدام نمایند .

– ایجاد شفافیت

- بکارگیری تجارب به منظور تشخیص نیازها، ارایه تنوع و بهبود عملکرد
  - بخش بندی جامعه هدف به منظور ارایه سرویس های کاملا سفارشی
  - جایگزینی تصمیم گیری انسانی با الگوریتم های اتوماتیک
  - نوآوری جدید در خصوص مدل های کسب و کار ، محصولات و سرویس ها
- استفاده از داده های عظیم به یکی از مزایای کلیدی رقابت و رشد سازمان ها تبدیل شده است : استفاده از داده های عظیم به یکی از روش های کلیدی سازمان های پیش رو برای عملکرد بهتر نسبت به رقبا تبدیل شده است . افزایش سهم خود در یک بازار رقابتی نمونه ای در این زمینه است . در دنیای رقابتی داده های عظیم ، یک بنگاه کسب و کار که نتواند ظرفیت های خود را به منظور استفاده از داده های عظیم ارتقاء دهد ، قطعاً نسبت به سایر رقبا عقب خواهد افتاد . داده های عظیم همچنین به ایجاد و رشد فرصت های جدید کمک خواهد کرد و طبیعی است که شاهد گروه جدیدی از شرکت ها باشیم که بطور کاملاً تخصصی بر روی موضوعاتی نظیر تجمیع و تحلیل داده فعالیت خواهند کرد . بسیاری از این بنگاه های اقتصادی ، شرکت هایی می باشند که بر روی حجم بالایی از جریان اطلاعات ( به عنوان نمونه : داده محصولات و سرویس ها ، خریداران و فروشندگان ، اولویت های مصرف کنندگان ) نشسته و قادر به ضبط و تحلیل داده خواهند بود . این موضوع برای شرکت هایی که دارای اینترفیس های مناسب با حجم بالایی از مشتریان خریدار یک مجموعه از محصولات و سرویس ها، شرکت های فعال در زنجیره تامین جهانی ، شرکت هایی که روزانه میلیون ها تراکنش را پردازش می کنند و شرکت هایی که پلت فرم های لازم برای تجربه دیجیتالی مصرف کنندگان را فراهم می کنند ، بسیار جذاب و قابل تامل است . تمامی این فرصت ها ، با مدیریت مناسب داده های عظیم فراهم می گردد . برخی از این فرصت ها ، منابع جدیدی از ارزش را تولید خواهند کرد و برخی دیگر ، باعث تغییر نگرش تولید ارزش در بسیاری از صنایع خواهند شد.
  - استفاده از داده های عظیم ، زیربنای موج جدیدی از رشد بهره وری و افزایش مصرف کننده را به دنبال خواهد داشت : بر اساس مطالعات انجام شده در پنج حوزه مختلف ، داده های عظیم باعث تقویت زیربنای مورد نیاز به منظور افزایش بهره وری می گردند . بگونه ای که سازمان ها را قادر می سازد خروجی با کیفیت بهتری را تولید کنند ( به عنوان نمونه افزایش ارزش افزوده محصولات و سرویس ها ) . مثلاً سازمان ها می توانند با بهره برداری از داده ، محصولاتی را طراحی نمایند که با نیاز مشتریان بهتر مطابقت می نماید . حتی می توان از داده در جهت بهبود محصولاتی که توسط مشتریان استفاده می گردد نیز کمک گرفت . به عنوان نمونه یک دستگاه موبایل را در نظر بگیرید که بر اساس شناختی که نسبت به رفتار صاحب خود پیدا کرده است ، می تواند برنامه ها و داده ها را متناسب با نیاز کاربر نگهداری نماید . این چنین

دستگاه های موبایل ، دارای ارزش بمراتب بیشتری نسبت به یک دستگاه موبایل جدید می باشند که بر اساس نیاز کاربر سفارشی نشده است.

- با این که استفاده از داده های عظیم، یک دغدغه مهم در بین حوزه های مختلف کسب و کار است ولی برخی از حوزه ها دارای استعداد بمراتب بیشتری جهت بهره برداری از مزایای آن می باشند: مطالعات انجام شده بین حوزه های مختلف کسب و کار نشان می دهد که الگوی بهره برداری از مزایای داده های عظیم بین آنها متفاوت است . حوزه اطلاعات و محصولات کامپیوتر و الکترونیک یکی از حوزه های پیشگام در این زمینه است که توانسته است با مدیریت صحیح داده های عظیم ، بهره وری خود را بطرز محسوسی افزایش دهد. بخش های مالی ، بیمه و دولتی حوزه بعدی استفاده از داده های عظیم به منظور افزایش کارایی می باشند . با این که تمامی حوزه ها لازم است که بر مشکلات موجود به منظور خلق ارزش با استفاده از داده های عظیم غلبه نمایند ، ولی مشکلات به سهم مساوی بین آنها تقسیم نشده است و برخی حوزه ها دارای مشکلات بمراتب بیشتری نسبت به سایر بخش ها می باشند . مثلا بخش های عمومی نظیر مراکز آموزشی با موانع جدی تری مواجه هستند . بدیهی است بخش هایی که تاکنون سرمایه گذاری اندکی بر روی فناوری اطلاعات کرده اند ، دارای مشکلات و موانع بیشتری می باشند .

- کمبود نیروی انسانی مستعد و ماهر در سازمان ها جهت بهره برداری از مزایای داده های عظیم : یکی از محدودیت های مهم به منظور بهره برداری از مزایای داده های عظیم و خلق ارزش ، کمبود نیروی انسانی مستعد و ماهر خصوصا در مواردی همچون حوزه آمار و یادگیری ماشین است. همچنین ، کمبود مدیران و تحلیل گرانی که با نحوه تولید بینش از داده های عظیم در سازمان ها آشنا باشند مزیت بر علت است . این انتظار وجود دارد که داده های عظیم به یکی از عوامل مهم و کلیدی رقابت بین بخش های مختلف تبدیل شود و تقاضا برای مشاغل تحلیلی در حوزه داده های عظیم روز به روز بیشتر گردد . هم اینک در اکثر قریب به اتفاق کشورهای جهان این نیاز با وضعیت موجود نیروی های مستعد در این زمینه همخوانی ندارد و تربیت این نوع نیروی انسانی مستعد ، کار ساده ای نیست و مستلزم سال ها آموزش است . به عنوان نمونه ، تربیت تحلیل گر داده های عظیم در امریکا تا سال ۲۰۱۸ به حدود ۳۰۰ هزار نفر خواهد رسید در صورتی که تا سال ۲۰۱۸ به چیزی حدود ۴۴۰ تا ۴۹۰ هزار نیروی انسانی مستعد در این حوزه نیاز است . با استناد به وضعیت بازار عرضه و تقاضای تحلیل گران داده در امریکا ، شاهد کمبود ۱۴۰ تا ۱۹۰ هزار نیروی انسانی مستعد در این حوزه می باشیم . این وضعیت صرفا در رابطه با امریکا صدق نمی کند و پیش بینی می شود که سایر کشورها خصوصا کشورهای صنعتی نیز وضعیتی مشابه داشته باشند . علاوه بر این ، پیش بینی شده است که در امریکا به ۵ / ۱ میلیون مدیر و تحلیل گر دیگر تا سال ۲۰۱۸ نیاز باشد که قادر به طرح صحیح سوالات و استفاده از نتایج تحلیل داده های عظیم بطور موثر باشند.

- به منظور بهره برداری از ظرفیت داده های عظیم به مسائل مختلفی باید پاسخ داده شود :

- سیاست های داده : به موازات افزایش حجم داده و جابجایی آن بین محدوده های سازمانی، به مجموعه ای از مسائل مرتبط با سیاست گذاری مواجه خواهیم شد که بسیارحائز اهمیت می باشند . حفظ حریم خصوصی، امنیت، مالکیت معنوی ومسئولیت نمونه هایی در این زمینه می باشند. حفظ حریم خصوصی بسیار حائز اهمیت است خصوصا در ارتباط با مصرف کنندگان . اطلاعات شخصی نظیر رکوردهای مالی و یا پرونده پزشکی افراد نمونه هایی در این زمینه می باشند . همچنین، امنیت داده نیز بسیار حائز اهمیت است . چگونه اطلاعات حساس و سایر اطلاعات مهم حفاظت و نگهداری گردند. اقتصاد رو به رشد داده های عظیم مجموعه ای از مسائل قانونی را به دنبال داشته است خصوصا اگر به این واقعیت توجه داشته باشیم که داده دارای تفاوت های اساسی نسبت به سایر سرمایه ها است .داده را می توان بطور کامل و به سادگی تکثیر و با سایر داده ها ترکیب نمود . بخشی از داده می تواند بطور همزمان و توسط بیش از یک فرد استفاده گردد . تمامی این موارد ، ویژگی های منحصر بفرد داده را در مقایسه با سایر سرمایه های فیزیکی نشان می دهد. در این خصوص لازم است به پرسش های مختلفی در رابطه با داده پاسخ داده شود : چه کسی مالک بخشی از داده است و چه حقوقی مرتبط با یک مجموعه داده است؟ استفاده منصفانه از داده چگونه تعریف می گردد ؟ همچنین، پرسش هایی در خصوص مسئولیت داده مطرح می شود: چه کسی مسئول داده غلطی می باشد که ممکن است مسائل و مشکلات بعدی را ایجاد نماید ؟

- فناوری و روش ها : به منظور استخراج ارزش از داده های عظیم، لازم است سازمان ها از فناوری های مختلف جدیدی ( به عنوان نمونه در حوزه هایی نظیر ذخیره سازی، محاسبات و نرم افزارهای تحلیلی) و روش های مدرن (نظیر انواع جدید تحلیل) استفاده نمایند. دامنه چالش های فناوری و مجموعه اولویت ها جهت برخورد با مسائل داده های عظیم با توجه به سطح بلوغ داده سازمان ها متفاوت است. سیستم های قدیمی، استانداردهای ناسازگار و قالب های نامتجانس، اغلب باعث بروزمشکلات عدیده ای در خصوص یکپارچگی داده و تجزیه و تحلیل های پیچیده می گردد. مسائل جدید و رشد قدرت پردازش ما را به سمت ایجاد روش های تحلیل جدید هدایت می کند . همچنین به نوآوری مستمر در حوزه فناوری ها و ابداع روش های جدیدی نیاز است تا به افراد و سازمان ها کمک کند تا بتوانند یکپارچه سازی، تحلیل، بصری سازی و مصرف داده های عظیم را به خوبی انجام دهند.

- تغییرات سازمانی و استعداد: رهبران سازمان ها اغلب دارای شناخت مناسبی از ارزش داده های عظیم نمی باشند و نمی دانند که چگونه می بایست از این فرصت طلایی در سازمان خود رمزگشایی نمایند. علاوه بر این، بسیاری از سازمان ها دارای نیروی انسانی مستعد جهت استخراج بینش از داده نمی باشند. همچنین، بسیاری از سازمان ها دارای ساختار مناسبی برای جریان های کاری و انگیزه های لازم برای بهینه سازی استفاده از داده های عظیم درجهت اتخاذ تصمیم بهتر نمی باشند.

- دستیابی به داده : برای فعال کردن فرصت تحول ، سازمان ها نیاز به یکپارچه سازی اطلاعات از چندین منبع داده دارند. در برخی مواردی ممکن است بتوان با پرداخت هزینه های مربوطه، دستیابی به یک منبع داده خاص را فراهم کرد ولی در بسیاری از موارد همچنان دستیابی به منابع داده خارج از سازمان مشکل و چالش های مختص به خود را دارد و در بسیاری از موارد امکان به اشتراک گذاشتن آن وجود ندارد چراکه ممکن است این تصور وجود داشته باشد که از داده موجود می توان به عنوان یک مزیت رقابتی برای کسب و کار خود استفاده کرد و نباید آن را به اشتراک گذاشت.

رشد سریع داده ها در بخش کلان داده ها طوفانی به پا کرد که چالش های بزرگی در اکتساب ، ذخیره سازی ، مدیریت و تحلیل داده ها به وجود آورد. سیستم های تحلیلی و مدیریت داده ی سنتی، مبتنی بر سیستم مدیریت پایگاه داده های رابطه ای<sup>۱</sup> بودند. چنین سیستم هایی فقط برای داده های ساخت یافته به کار گرفته می شوند نه داده های نیمه ساخت یافته و غیر ساخت یافته. همچنین ، استفاده ی افزایشی از RDBMS باعث استفاده از سخت افزارهایی باقیمت رو به افزایش می شود. ظاهراً RDBMS های سنتی قادر به اداره ی این حجم بزرگ و ناهمگن کلان داده ها نیستند. جامعه ی تحقیقاتی راه کارهایی از جنبه های مختلف پیشنهاد کرده است. به عنوان مثال برای نیازمندی های زیرساخت کلان داده ها مثل بهره وری هزینه، قابلیت کشسانی<sup>۲</sup> و کم و زیاد کردن بدون اشکال منابع از رایانش ابری استفاده شد. سیستم فایل های توزیع شده و پایگاه داده های NoSQL برای ذخیره سازی پایدار و مدیریت مجموعه داده های نامرتب مقیاس - بزرگ ، انتخاب های خوبی بودند. چنین چارچوب های برنامه نویسی در پردازش وظایف<sup>۳</sup> خوشه بندی شده، به ویژه برای رتبه بندی صفحات وب ، بسیار موفق بودند(هووارد، ۱۹۹۸). کاربردهای مختلف کلان داده ها می توانند بر اساس این فناوری ها و سکو های<sup>۴</sup> نوپا توسعه یابند و این موضوع برای توسعه سیستم های تحلیل کلان داده ها بی اهمیت نیست. مقالات بسیاری راجع به موانع سر راه توسعه ی کاربرد های کلان داده ها بحث کرده اند(کاتل، ۲۰۱۱). بعضی چالش های کلیدی در ادامه آمده است:

## ۲-۲-۵ : نمایش داده<sup>۵</sup>

بسیاری از مجموع داده ها سطوح معینی از ناهمگنی در نوع ، ساختار، معناشناسی<sup>۶</sup>، سازمان دهی ، دانه بندی<sup>۷</sup> و قابلیت دسترسی دارند. هدف نمایش داده بیشتر معنادار نمودن داده ها برای تحلیل کامپیوتری و تفسیر توسط کاربران است. یک نمایش داده ی نامناسب ارزش داده ی اصلی را کاهش خواهد داد و حتی ممکن است مانع تحلیل موثر داده شود. نمایش

---

<sup>۱</sup>RDBMS

<sup>۲</sup>Elasticity

<sup>۳</sup>Tasks

<sup>۴</sup>Platforms

<sup>۵</sup>Data Representation

<sup>۶</sup>Semantics

<sup>۷</sup>Granularity



داده ی کارا باید منعکس کننده ساختار ، کلاس، نوع و همچنین فناوری های یکپارچه<sup>۱</sup> داده ها باشد تا بتواند عملگرها های موثر را روی مجموعه داده های متفاوت به کار گیرد(لابریدینیس و جاگادیش، ۲۰۱۲).

## ۲-۲-۶ فشردگی و کاهش افزونگی داده ها<sup>۲</sup>

معمولا سطح بالایی از افزونگی در مجموعه داده ها وجود دارد. فشردگی و کاهش افزونگی داده ها برای کاهش غیر مستقیم هزینه ی کلی سیستم موثر است با این فرض که ارزش های بلقوه ی آن داده ها تحت تاثیر این دو عمل واقع نشود. به عنوان مثال، اغلب داده های تولید شده توسط شبکه های حسگر دارای افزونگی بالایی هستند که ممکن است بر اساس درجه ی اهمیت، فیلتر و فشرده شده باشند.

## ۲-۲-۷ مدیریت چرخه ی زندگی داده ها<sup>۳</sup>

در مقایسه با پیشرفت نسبتا کند سیستم های ذخیره سازی اما محاسبات و حسگر های فراگیر با نرخ و مقیاسی بی نظیر در حال تولید داده ها هستند. ما با چالش های مهم زیادی روبرو هستیم از جمله این که سیستم های ذخیره سازی فعلی چنین داده های انبوهی را پشتیبانی نمی کند. در واقع ، ارزش های پنهان شده در کلان داده ها به تازه بودن داده ها<sup>۴</sup> بستگی دارد. بنابراین باید یک اصل مهم وابسته به ارزش تحلیلی توسعه یابد که مشخص کند چه داده ای باید ذخیره شود و چه داده ای باید دور ریخته شود.

## ۲-۲-۷ مکانسیم تحلیلی<sup>۵</sup>

سیستم تحلیلی کلان داده ها باید انبوهی از داده های ناهمگن را در مدت زمان محدودی پردازش کند. اما RDBMS سنتی با طراحی انعطاف ناپذیر و فقدان مقیاس پذیری و توسعه پذیری که دارد برای برآوردن نیازمندی های کارایی مناسب نیست. پایگاه داده های غیر رابطه ای<sup>۶</sup> مزایای منحصر به فردی در پردازش داده های غیر ساخت یافته از خود نشان داده بودند و به عنوان مسیر اصلی در تحلیل کلان داده ها آغاز به کار کرده بودند. با این وجود، در پایگاه داده های غیر رابطه ای هنوز برخی مشکلات در کارایی و کاربرد های خاص آن ها وجود دارد. ما باید به دنبال یک راهکار سازشی بین RDBMS ها و پایگاه داده های غیر رابطه ای باشیم . به عنوان مثال ، بعضی موسسات از معماری پایگاه داده ترکیبی استفاده کرده اند که مزایای هر دو نوع پایگاه داده را مجتمع می کرد (مثل فیس بوک ). در زمینه ی پایگاه داده های in-memory و الگوی داده ی مبتنی بر تحلیل تقریبی به تحقیقات بیشتری نیاز است.

<sup>۱</sup>Integrated technologies

<sup>۲</sup>Redundancy Reduction and data Compression

<sup>۳</sup>Data Life Cycle Management

<sup>۴</sup>Data freshness

<sup>۵</sup>Analytical Mechanism

<sup>۶</sup>Non-relational databases

## ۸-۲-۲ محرمانگی داده ها

در حال حاضر اکثر مالکان یا فرآهم آورندگان سرویس های کلان داده ها به دلیل توانایی محدودشان نمی توانند چنین مجموعه ی داده های انبوهی را به صورت موثر نگهداری و تحلیل کنند. آن ها برای تحلیل بیشتر داده ها به ابزارها یا تخصص هایی تکیه می کنند که باعث افزایش خطرات امنیتی بالقوه می شود. به عنوان مثال، مجموعه داده های تراکنشی معمولاً شامل داده های عملیاتی کاملی هستند تا پردازش های تجاری را پیش ببرند. این داده ها شامل جزئیاتی از کوچیک ترین دانه بندی و اغلب اطلاعات حساسی از قبیل شماره های کارت اعتباری هستند. بنابراین، برای تامین امنیت، فقط زمانی تحلیل کلان داده ها ممکن است به شخص ثالثی<sup>۱</sup> واگذار شود که اقدامات امنیتی مناسبی برای حفاظت از داده های حساس صورت گرفته باشد (آگراوال و همکاران، ۲۰۱۲).

## ۹-۲-۲ مدیریت انرژی

مصرف انرژی در سیستم های محاسباتی کامپیوترهای بزرگ هم از لحاظ اقتصادی و هم از لحاظ محیط زیست مورد توجه است. با افزایش حجم و نیازمندی های تحلیلی داده ها، پردازش، ذخیره سازی و انتقال کلان داده ها ناچار به مصرف بیشتر و بیشتر انرژی الکتریکی هستیم بنابراین برای تامین قابلیت توسعه و قابلیت دسترسی کلان داده ها باید مکانسیم های مدیریتی و کنترل مصرف انرژی سیستم به کار برده شود.

## ۱۰-۲-۲ توسعه پذیری و مقیاس پذیری<sup>۲</sup>

سیستم تحلیلی کلان داده ها باید مجموعه داده های فعلی و آینده را پشتیبانی کند. الگوریتم های تحلیلی بایدتوانایی پردازش مجموعه داده های پیچیده تر و به صورت فزاینده در حال توسعه را داشته باشند.

## ۱۱-۲-۲ همکاری<sup>۳</sup>

تحلیل کلان داده ها یک تحقیق میان رشته ای است که برای ثمر دادن پتانسیل کلان داده ها نیاز به همکاری کارشناسان در زمینه های مختلف دارد. باید یک معماری شبکه ای کلان داده های جامع ایجاد شود تا به دانشمندان ومهندسان برای دسترسی به انواع مختلف داده و استفاده کامل از توانایی هایشان کمک کند تا برای اهداف تحلیلی با یکدیگر همکاری کنند (آگراوال و همکاران، ۲۰۱۲).

---

<sup>۱</sup>Third party

<sup>۲</sup>Expendability and Scalability

<sup>۳</sup>Cooperation

## ۲-۲-۱۲ پیدایش ارزش نظام مند کسب و کار

داده‌های بزرگ نوین، که اغلب داده‌های چندساختاری خوانده میشود، فرصت هایی را در زمینه بهینه سازی فرآیندهای عملیاتی و ابداع فرآیندهای نوین و همچنین در علم تحلیل و هوش کسب و کار به وجود آورده است. این فرصت ها را می توان در قالب چهار گونه گسترده از کارکردهای کسب و کار که به وسیله داده‌های بزرگ توانمند شده اند دسته بندی کرد:

### ۱. توسعه مدل کسب و کار با استفاده از داده‌های بزرگ

ایجاد درآمد و توسعه مدل کسب و کار، به خصوص در صنعت خرده فروشی و کالاهای مصرفی که در آن یک تعامل گسترده مستقیم بین سازمان ها و بازارهای مصرفی بزرگ وجود دارد، از طریق دسترسی به داده‌های بزرگ بسیار آسان می شود. بخش بازاریابی در شرکت های پیشرو از اطلاعات رسانه های اجتماعی هم از جهت محتوایی و هم از جهت رابطه ای استفاده می کنند و رویکرد خود را از نمونه گیری به سمت دربرگرفتن تمامی داده‌ها تغییر داده اند. آنها شیوه بخش بندی بازار خود را تغییر داده و از تحلیل درازمدت روندهای داده‌های تاریخی عدول کرده و روی به واکنش فوری به رخداد های تازه آورده اند. پیش بینی رفتار مشتریان و نتایج اقدامات پیشنهادی اجازه خلق شدن و آزمون شدن را به مدل های تازه کسب و کار می دهد.

### ۲. مدیریت قیمت ها به صورت فوری با استفاده از داده‌های بزرگ

با کمک داده‌های بزرگ امکان پایش رخدادها به صورت همراه در گوشی های همراه فراهم شده و پی بردن به کلاهبرداری ها در داده‌های تراکنش های مالی آسان تر شده است و زمان کمتری نسبت به گذشته می برد. استفاده از فنون تحلیل داده‌های بزرگ و پی بردن به جریان داده‌ها پیش از ذخیره سازی آنها به یک هنجار تبدیل شده و واکنش سریع به مشکلات خاص را پیش از اینکه اوج بگیرند و گسترش یابند فراهم می کند.

### ۳. پیش بینی همزمان

شرکت های پیشرو با استفاده از فناوری های پردازش داده‌های بزرگ و با استفاده از سنسورهای پیشرفته در وسایلی همچون تامین برق و آب و تجهیزات ارتباطات از راه دور، از اندازه گیری های کلی و کلان به سمت اندازه گیری های خرد و دقیق رو آورده اند. این امر امکان پیش بینی روندهای خرد را فراهم می آورد و با افزایش مصرف باعث افزایش ارزش کسب شده به وسیله این گونه شرکت ها می شود.

### ۴. بازآفرینی فرآیندهای کسب و کار

داده‌های بزرگ از جهت استفاده نوآورانه از داده‌های خلق شده به وسیله حسگرهای دقیق، امکان بازآفرینی همه صنایع را فراهم آورده است. مثلا در بیمه خودروها می توان بیمه نامه هایی را بر مبنای رفتار واقعی مشتریان و نه بر اساس میانگین

های آماری از مخاطرات و ریسک ها تنظیم نمود. دسترس پذیری داده‌های ژنتیکی مربوط به افراد و سوابق الکترونیکی پزشکی افراد فرصت های چشمگیری را برای صنایع بیمه پیش آورده است. هر چند که این مساله از نظر اخلاقی مورد اختلاف است.

#### ۵. ارائه راه حل های هوش تجاری

تعریف شاخص های مناسب در هر حوزه کسب و کار ابتدای راه تعریف راه حل های هوش تجاری است. این راه حل ها تعاریف موجود برای اصطلاحات کلیدی و شاخص های محیط هوش تجاری را مستند کرده و قوانین موجود را پیاده سازی می کنند. مستند سازی شاخص های کلیدی عملکردی (KPI) سازمان توسط واحد هوش تجاری با کمک تمامی واحدهای کسب و کار که قرار است در راه حل هوش تجاری سازمان به کار روند انجام میشود. برای استفاده از این راه حل ها نیازمند داشتن زیر بنای اطلاعاتی هستیم. با معماری صحیح زیر ساخت ها و شناخت دقیق روابط بین کلان داده ها به ما در ارائه وضعیت موجود در هر حوزه و اخذ تصمیم های مناسب برای بهبود و ارتقا حوزه کسب و کار، یاری میرساند. در واقع کلان داده ها در ساخت ابزار تصمیم یار نقش پایه ای دارند.

#### ۲-۲-۱۱ کشف دانش از داده های حجیم (KDD)

کشف دانش از داده های حجیم عبارت است از تعدادی عملگر که برای دریافت اطلاعات از مجموعه های داده پیچیده ، طراحی شده اند. رئوس کلی اکتشاف دانش از داده های حجیم:

۱. دامنه برنامه کاربردی مقدم به اطلاعات و تعریف هدف از پردازش از دیدگاه مشتری.
۲. ایجاد زیر مجموعه ای از داده ها که به کشف دانش اشاره میکنند.
۳. از بین بردن noise ها، اداره کردن فیلدهای داده گمشده ، جمع آوری اطلاعات مورد نیاز جهت مدلسازی و محاسبه زمان اطلاعات و تغییرات اصلی.
۴. پیدا کردن خصوصیات مفید جهت نشان دادن وابستگی داده به هدف از کار.
۵. نگاشت اهداف به روشهای داده کاوی دقیق.
۶. انتخاب الگوریتم داده کاوی و تابع جهت جستجوی الگوهای داده.
۷. جستجوی الگوها به شکل گویا.
۸. بازگشت به هر مرحله ۱ تا ۷ جهت تکرار و بازگویی ، همچنین این مرحله میتواند شامل تصویرسازی از الگوها باشد.

۹. استفاده از اطلاعات به شکل مستقیم ، ترکیب اطلاعات در سیستم دیگر یا گزارش گیری و استفاده از اطلاعات به شکل ساده.

۱۰. کشف دانش از داده های حجیم با استفاده از Hadoop را در ۳ اصل تحلیل می شود که عبارتند از: [۱]

۱- کشف دانش از داده ها ، شامل روشهای تحلیلی متنوعی میباشد مانند برنامه نویسی توزیع شده، الگوشناسی، داده کاوی، پردازش زبان طبیعی، تجزیه و تحلیل احساسات، تجزیه و تحلیل آماری و تصویری و تعاملات کامپیوتری بشر. بنابراین معماری باید روشهای تحلیل و متدهای متنوعی را پشتیبانی کند.

- تجزیه و تحلیل آماری ، علاقه مند در خلاصه سازی مجموعه داده های حجیم ، فهم داده ها و تعریف مدلهایی جهت پیشگویی میباشد.

- داده کاوی با کشف الگوهای مفید در مجموعه داده های حجیم از طریق خودشان در ارتباط است، آموخته های ماشین که با داده کاوی ترکیب میشود و روشهای آماری که ماشینها را قادر به فهم مجموعه های داده میسازد.

- روشهای تجزیه و تحلیل بصری ، ناحیه ای را که در آن مجموعه داده های حجیم به کاربران سرویس میدهند را توسعه میدهند ، در راه های چالش برانگیز قادر به فهم روابط خواهند بود.

۲- یک معماری KDD جامع باید نگهداری و به کار انداختن خط پردازش را تأمین کند.

- آماده سازی داده ها تجزیه و تحلیل دسته ای جهت عیب یابی مناسب error ها، مقادیر از دست رفته و فرمت های غیر قابل استفاده.

- پردازش داده های ساخت یافته و داده های نیمه ساخت یافته.

۳- این یک اصل است که نتایج ساده و در دسترس باشند. به همین دلیل، روشهای زیر جهت رسیدن به این موضوع ، استفاده میشوند:

- استفاده از استانداردهای رایج و متن باز.

- استفاده از معماری های مبتنی بر web (لابریدینیس و جاگادیش، ۲۰۱۲)

## ۲-۳-۱ تأثیر big data بر فرآیند بهبود تحلیل کسب و کار

با ورود به عصر اطلاعات و ارتباطات و آغاز استفاده از داده ها و اطلاعات به عنوان سرمایه های اصلی در حرکت علمی، اقتصادی، اجتماعی و فرهنگی جوامع، سازمان ها و شرکت های مختلف و توسعه مشارکت افراد در جهان اینترنت و ارتباطات شبکه ای در دنیا، دغدغه ای بروز پیدا کرد که از جنس همین داده هایی بود که همه روز و با سرعت وحشتناک در دنیا و در عرصه های مختلفی که فناوری اطلاعات ورود پیدا کرده بود، تولید می شود و آن اینکه چگونه این حجم بزرگ و متنوع داده ها و اطلاعات را با توجه به ساختار هایی که در فضای فناوری اطلاعات وجود دارد، می توان مدیریت، کنترل و پردازش کرد و از آن در جهت بهبود ساختارها و سودآوری بیشتر بهره جست؟. تحلیل مجموع داده ها می تواند همبستگی های جدید را پیدا کند، که مربوط به روندهای تجارت نقطه ای، جلوگیری از بیماری، جرایم جنگی و ... می باشد. دانشمندان، روسای تجاری، کارورزان صدا و سیما، تبلیغات و دولت، همانند مقررات، باعث مشکلاتی برای کلان داده در حوزه های مانند جستجوی اینترنتی، مالی و اطلاعات تجاری می شوند. دانشمندان با محدودیت هایی در کارهای علوم مواجه اند که شامل هواشناسی، ژنتیک و اطلاعات، شبیه سازی فیزیکی پیچیده، و تحقیقات محیط زیستی و بیولوژیکی می باشد.

اندازه مجموعه اطلاعات در بخشی افزایش می یابد، به دلیل اینکه آنها از طریق ابزار موبایل که ارزان و بیش شمارند، آنتن هوایی (دریافت از راه دور) وقفه نرم افزاری، دوربین ها، میکروفن ها، خواننده فرکانس رادیویی مشخص، و شبکه های دریافت بی سیم، جمع آوری می شوند. ظرفیت جهانی تکنولوژی در هر واحد سرمایه، به منظور جمع آوری اطلاعات از دهه ۱۹۸۰ هر ۴۰ ماه دو برابر می شود. از سال ۲۰۱۲ هر روزه ۲۰۵ اگزابایت (۱۰۱۸×۲۰۵) اطلاعات ایجاد شده است. چالش موسسات بزرگ این است تعیین کنند که چه کسی ابتکار عملیات کلان داده را دارا می باشد که کل سازمان را در بر می گیرد.

سیستم های مدیریت پایگاه داده رابطه ای، و وضعیت صفحه نمایش و بسته تصویری اغلب با مشکلاتی در رابطه با دستکاری در کلان داده مواجه اند. در عوض اینکار به اجرای نرم افزاری به طور همزمان در حجم گسترده و با ده ها، صدها، و حتی هزاران سرور، نیاز دارد. آن چیزی که به عنوان کلان داده در نظر گرفته می شود، بر مبنای توانایی های استفاده کنندگان از وسایل تفاوت دارد و گسترش توانایی ها برای ایجاد کلان داده یک هدف می باشد؛ بنابراین، هر آنچه که کلان در نظر گرفته می شود یک سال بعد عادی می شود. برای برخی سازمانها، که با صدها گیگابایت اطلاعات برای اولین بار مواجه اند، ممکن به در نظر گرفتن دوباره اختیارات مدیریت اطلاعات نیاز احساس شود. برای دیگران، ممکن است ده ها یا صدها ترابایت باید حجم اطلاعات افزایش یابد تا به عنوان قابل توجه در نظر گرفته شود. اکثر شرکتها با مقدار زیادی داده های جدید که در اشکال مختلف بسیاری دریافت میشوند، روبرو هستند. کلان داده ها این پتانسیل را دارد که بینش هایی ایجاد کند که میتوانند هر کسب و کاری را دگرگون کند.

کلان داده (Big Data) وعده‌های زیادی به صنعت بازاریابی داده است. اصلی‌ترین وعده، پاسخ به دو سوال اساسی است. اول اینکه با کمک آن‌ها بتوان فهمید چه کسی، چه محصولی را در چه زمانی با چه قیمتی می‌خرد؟ و دوم اینکه چگونه می‌توان بین محصولی که مشتریان می‌بینند یا در موردش شنیده یا می‌خوانند، و چیزی که نهایتاً می‌خرند یا مصرف می‌کنند، ارتباطی منطقی پیدا کرد؟

پاسخ به این دو سوال، بازاریابی را بسیار اثرگذارتر از روش‌هایی مثل هدفگیری و شناسایی مشتری می‌کند؛

۱. روش‌هایی که نیمی از بودجه بازاریابی را هدر می‌دهند.

بازاریاب‌ها برای یافتن پاسخ باید آموزش ببینند تا بتوانند از کلان داده‌ها در این مسیر و برای پیش‌بینی خرید بعدی مشتریان، بهره ببرند. در این راه، بازاریاب‌ها باید تلاش کنند در تشخیص الگوی مصرف هر مشتری و شناخت علاقمندی‌ها، خواسته‌ها و عادات خرید و مصرف او، موفق و بهینه عمل کنند. نتیجه، جزیاتی دقیق و کامل از عملکرد مصرف‌کننده‌ها و امکان پیش‌بینی خریدهای بعدی آن‌ها خواهد بود.

ما بسیاری از صنایع در تلاش برای تشخیص سریع‌تر اهداف مشتری‌ها، با حقیقتی نگران‌کننده روبرو شده‌اند. تکنیک‌های مورد نظر برای پیش‌بینی خریدها و جذب مشتریان، تنها برای بازه‌های کوتاه مدت نتیجه بخش است. در این بازی، رقابت بازاریاب‌ها برای شناسایی حرکت بعدی مشتریان به توازن رسیده و تلاش‌ها بر کسب سود حاشیه‌ای متمرکز شده‌اند. نتیجه، رقابتی بی‌برنده در بازه‌ی زمانی کوتاه مدت است. همه در بازه‌های زمانی میان مدت و بلندمدت دوشادوش هم در حرکت به سوی برابری پیش رفته و هیچ‌یک از مدعیان مزیت رقابتی پایداری برای پیروزی در اختیار ندارند.

البته معنی گفته‌های بالا، توقف تلاش شرکت‌ها برای پیش‌بینی بهینه‌تر رفتار مشتریان نیست؛ اما زمانی که همه‌ی رقبا مشغول همین تلاش هستند و در شرایطی که ماندن در بازی هم پاداش‌های خودش را برای صنایع مختلف دارد، حداکثر نتیجه‌ای که از این تلاش به دست می‌آید، کسب درآمد بالاتر از حد میانگین خواهد بود. در بسیاری صنایع از جمله گردشگری، بیمه، ارتباطات، موسیقی و حتی خودروسازی، تکنیک‌های بازاریابی به حدی پیشرفت کرده که رقبا به سطحی برابر در قدرت جذب مشتری رسیده‌اند. مهم‌ترین و تنها مولفه‌ی باقیمانده برای جذب مشتری بیشتر، قدرت پیش‌بینی رفتار مصرف‌کننده‌های نهایی است.

برای دستیابی به این مزیت رقابتی، برنامه‌های بازاریابی، به تحلیل اطلاعات کلان داده‌ها برای پاسخ به سوالات استراتژیک درباره‌ی چگونگی جذب بیشتر مشتریان و ایجاد حس وفاداری و برقراری ارتباط پایدار با آن‌ها، نیاز دارند؛ ارتباطی که بتواند با پایداری در بلندمدت، نیازهای مشتریان را در مقابل رقبا، مرتفع کند.

سوال‌های استراتژیک یادشده، به موارد مهمی همچون کشف روش‌های ایجاد حس وفاداری در مشتریان کمک شایانی می‌کنند. در شرایط فعلی، قیمت‌گذاری مناسب برای جذب بیشتر مشتری، کارساز نیست، بلکه باید راهی برای ایجاد ارزش بیشتر در محصولات و خدمات پیدا کرد که مشتریان قدر آن را بدانند. یافتن روش‌هایی برای عدم رغبت مشتریان به امتحان کردن یا انتخاب محصول و خدمات رقیب، استراتژی‌ای بسیار مهم‌تر از تبدیل مشتریان رقیب به مصرف‌کنندگان کالاها یا خدمات خود است.

پاسخ به این سوالات اساسی و یافتن راه‌حل‌های مطلوب، نیازمند استفاده‌ی متفاوت از کلان داده‌ها است. لازم است به جای یافتن راه‌حل‌هایی برای هدف‌گیری مشتریان بالقوه، به دنبال استفاده از کلان داده‌ها جهت کسب صفات ارزشمند در محصولات و خدمات بود. این صفات ارزشمند، چیزهایی هستند که مشتریان برای انتخاب گزینه‌ی ایده‌آل به آن توجه می‌کنند. سوال اساسی و متفاوت این است: کلان داده‌ها چگونه می‌توانند به ما و نه مشتریان کمک کنند؟

با کمک کلان داده‌ها می‌توان اطلاعاتی جدید را برای ایجاد محصولات و خدماتی کاملاً تازه کشف کرد. مثال ساده در همین رابطه، استفاده از کلان داده‌ها برای کمک به تسریع روند یافتن نتیجه‌ی مورد نظر کاربران به هنگام استفاده از موتورهای جست‌وجو باشد. روشی که در حال حاضر شرکت‌هایی مانند آمازون و نت‌فلیکس برای کمک به کاربران‌شان جهت یافتن سریع‌تر گزینه‌های مورد نظرشان، از آن بهره می‌برند. در مثالی دیگر می‌توان به تجمیع منابع اطلاعاتی‌ای که شرکت‌ها از نتیجه‌های جست‌وجوی سایر مشتریان و کاربران‌شان به‌دست می‌آورند اشاره کرد. اطلاعاتی که با کمک آن‌ها می‌توان با تشخیص روند جست‌وجوی سایر مشتریان و مقایسه‌ی آن با مشتری فعلی، نتایج بهتری را در اختیار او گذاشت.

نگاهی به استارت‌آپ‌هایی که امکان استخراج اطلاعات قابل اتکاتری از کلان داده‌ها فراهم کرده‌اند، می‌تواند جذاب باشد.

شرکت **Opower** به کاربران اجازه می‌دهد با به اشتراک‌گذاری صورت‌حساب‌هایشان با دوستان فیسبوکی خود، رتبه‌ی خود را در قیاس با دیگران مشاهده کنند. **INRIX** با تجمیع میزان ترافیک مصرفی کاربران، امکان گزارش‌گیری بی‌درنگ را مهیا کرده است. شرکت شناخته‌شده‌ی **Zillow** با ادغام اطلاعات از منابع مختلف، آماری تلفیقی از ویژگی‌ها، قیمت و سایر صفات موثر در معاملات املاک را در اختیار خریداران، فروشندگان و معامله‌گرهای ملک قرار می‌دهد.

این شرکت‌ها، استفاده از کلان داده‌ها را سنگ‌بنای کسب‌وکارشان قرار داده‌اند و مسیر موفقیت‌شان باید الگویی برای تمامی شرکت‌ها در حوزه‌های مختلف باشد. چرا که امروزه هیچ کسب‌وکاری نمی‌تواند بدون استفاده‌ی موثر از داده‌ها، به حیات خود ادامه دهد.

هر شرکت باید سه سوال مهم زیر را برای ماندن در بازار رقابت در نظر بگیرد:



۱. چه نوع اطلاعاتی می‌تواند هزینه و ریسک مشتریان نهایی را به حداقل برساند؟

شرکت‌های بزرگی مانند Yelp، Zagat، TripAdvisor، Uber، eBay، Netflix و Amazon اهمیت زیادی به جمع‌آوری اطلاعات مرتبط با رتبه‌بندی تامین‌کننده‌های کالا و خدمات می‌دهند. چرا که این اطلاعات میزان ریسک‌پذیری مشتریان را به حداقل می‌رساند. در حال حاضر این دست اطلاعات، فاکتور مهمی برای ارزیابی عملکرد تامین‌کننده‌ها به حساب می‌آیند. با این حال، مشتریان به طرز فزاینده‌ای به اطلاعاتی که کمک کند بتوانند رفتار کاربرانی مثل خودشان را در مقابله با محصولات و خدمات مختلف ببینند، توجه نشان می‌دهند. در اختیار گذاشتن چنین اطلاعاتی به قدرت درک عمیق رفتار مشتریان و کشف الگوی توجه‌شان به عملکرد یکدیگر نیاز دارد؛ و این چیزی است که با کمک کلان داده‌ها می‌توان به آن دست یافت.

۲. چه نوع اطلاعاتی در دسترس هستند؟ ۳. آیا باید آن‌ها را جمع‌آوری و مورد استفاده قرار داد؟

آیا جمع‌آوری اطلاعاتی حاشیه‌ای مانند موقعیت مکانی کاربران می‌تواند ارزشمند باشد؟ این چیزی است که InVenture که یک استارت‌آپ جذاب در آفریقا است، به آن پرداخته است. این شرکت با تبدیل اطلاعاتی که به صورت اتفاقی در تلفن همراه کاربران تولید می‌شود، امکان تشخیص مشتریانی را که به خدمات مالی خاصی دسترسی یا نیاز داشته یا دارند برقرار کرده است. در محیطی که هیچ تاریخچه‌ای از وضعیت اعتبار مالی بخش اعظمی از جمعیت وجود ندارد، استفاده از اطلاعات حاشیه‌ای که به ظاهر بی‌اهمیت به نظر می‌رسند، به دستیاری مناسب برای موسسات مالی در جهت تصمیم‌گیری برای پرداخت یا عدم پرداخت وام به مشتریان تبدیل شده است. برای مثال این شرکت تشخیص داده که احتمال بازپرداخت منظم و کامل وام افرادی که در دفترچه تماس تلفن همراه‌شان نام و نام‌خانوادگی افراد را به طور کامل ثبت کرده‌اند، بیشتر است.

تنوع و اختلاف سلیقه‌ی مشتریان، چگونه بر میزان سود حاصل از جمع‌بندی اطلاعات سایرین تاثیر می‌گذارد؟

برای مثال شرکتی که محصولات مرتبط با کشاورزی مانند بذر، کود و آفت‌کش را عرضه می‌کند، می‌تواند با جمع‌آوری اطلاعات خرید کشاورزان تشخیص دهد که هر کشاورز بر اساس نوع زمین و شرایطش، چه خرید بهینه‌ای را می‌تواند انجام دهد. جمع‌آوری اطلاعات تعداد زیادی کشاورز در کنار طبقه‌بندی مولفه‌هایی مانند نوع خاک، آب‌وهوا و شرایط زمین هر یک از آن‌ها، می‌تواند کمک کند که برای هر کشاورز، بسته‌ی بهینه‌ای از محصولات پیشنهاد شود و در نهایت نتیجه‌ی بهتری نیز برداشت شود.

کلان داده‌ها می‌توانند بسیاری از سوالات اساسی بازاریاب‌ها را پاسخ دهند. یافتن پاسخ سوالاتی که می‌تواند برای مدت‌ها برای آن‌ها بی‌جواب مانده باشد با کمک گرفتن و تحلیل صحیح کلان داده‌ها امکان‌پذیر می‌شود. این، تنها راه‌حل برای صنعت بازاریابی است که با کمک مجموعه‌ای از اطلاعات، بهترین بازخورد و بالاترین سطح بازدهی را در درازمدت برای صنایع مختلف، حاصل کند.

تأثیرات استفاده از کلان داده ها در شرکت های مختلف

## ۱. تجزیه و تحلیل و بهینه سازی شبکه

شرکت های ارایه دهنده سرویس های ارتباطی برای برخورد با حجم انفجاری داده موبایل ، نیازمند سرمایه گذاری گسترده ای بر روی شبکه خود می باشند و معمولاً بین ۱۸ تا ۲۰ درصد از درآمد هر سال خود را صرف CAPEX می کنند . ظرفیت شبکه ، یکی از منابع ارزشمند است و شرکت های تلکام به دنبال استفاده گسترده از کلان داده ها و تجزیه و تحلیل داده جهت مانیتورینگ موثر و مدیریت ظرفیت شبکه ، ایجاد مدل های پیش بینی ظرفیت ، اولویت بندی و تصمیم گیری در خصوص برنامه ریزی توسعه شبکه می باشند .

بهینه سازی و برنامه ریزی ظرفیت شبکه: با تلفیق استفاده از شبکه ، دانسیته مشترک به همراه داده مکان و ترافیک ، شرکت های ارایه دهنده سرویس های ارتباطی قادر به مانیتورینگ با دقت بیش تر ، مدیریت و پیش بینی ظرفیت شبکه و برنامه ریزی موثر برای قطعی احتمالی می باشند . با استفاده از داده وضعیت ظرفیت به صورت بلادرنگ ، شرکت های ارایه دهنده خدمات ارتباطی قادر به بصری سازی و مشاهده مناطق با تراکم بالا در مواردی می باشند که با توجه به میزان ظرفیت پیش بینی شده ، احتمال قطعی به شدت افزایش می یابد تا بتوانند اولویت بندی جهت توسعه ظرفیت جدید را انجام دهند. همچنین، برای مناطقی که دارای ظرفیت بیش از حدی می باشند ، می توان با اجرای کمپین های خاص و یا تبلیغات اقدام به جذب مشتری جدید و یا بسته های تشویقی مختص مشترکین موجود به منظور استفاده حداکثری از ظرفیت شبکه کرد.

بهینه سازی و بکارگیری موثر ظرفیت شبکه باعث صرفه جویی میلیون ها دلار برای شرکت های تلکام طی سال می گردد . به کمک تجزیه و تحلیل های بلادرنگ و میزان ترافیک ، شرکت های ارایه دهنده خدمات ارتباطی می توانند مدل های پیش بینی ظرفیت را ایجاد نمایند و با مانیتورینگ مستمر میزان واقعی ترافیک در مقابل ترافیک پیش بینی شده در گذشته ، از بینش به دست آمده جهت بهینه سازی مدل و برنامه ریزی برای ظرفیت مکمل در مواردی که احتمال قطعی وجود دارد ، استفاده نمایند.

## ۲. توسعه شبکه و برنامه ریزی جهت سرمایه گذاری :

برنامه ریزی و اولویت بندی پروژه های توسعه شبکه همواره برای شرکت های ارایه دهنده خدمات ارتباطی مشکل است . با توجه به وجود وابستگی ها و ملاحظات متعدد ، شرکت های ارایه دهنده خدمات ارتباطی می بایست دارای توانمندی لازم جهت برنامه ریزی موثر به منظور تعیین اولویت های سرمایه گذاری و منابع بر اساس نیازهای ارتباطی آینده ، اهداف استراتژیک ، برگشت سرمایه گذاری ( ROI ) پیش بینی شده ، ترافیک پیش بینی شده ، تجربه مشتری و ... باشند . در تمامی موارد فوق ، می بایست به این نکته مهم توجه شود که بیش ترین منفعت حاصل از سرمایه گذاری می بایست

نصب مشتریان با ارزش بالا گردد. شرکت های ارایه دهنده خدمات ارتباطی می بایست قادر به ترکیب داده ترافیک شبکه ، معیارهای تجربه مشتری، داده درآمدهای بالقوه و مکان به همراه داده ارزش مشتری باشند تا این اطمینان حاصل گردد که آنها سرمایه گذاری CAPEX را در نقاط درست انجام داده اند . تعدادی از شرکت های ارایه دهنده خدمات ارتباطی از هدوپ و ابزارهای تجزیه و تحلیل کلان داده ها به منظور کمک در توسعه شبکه و اهداف برنامه ریزی شده استفاده می کنند . به عنوان مثال شرکت BT از هدوپ و تجزیه و تحلیل کلان داده به منظور تشخیص این که چگونه و کجا خدمات پهن باند با سرعت بالا را در انگلستان توسعه دهد، استفاده می نماید .

#### ۱. تجزیه و تحلیل بلادرنگ شبکه :

شرکت های ارایه دهنده خدمات ارتباطی از کلان داده و ابزارهای تحلیل جهت ایجاد heat map ظرفیت بلادرنگ استفاده می نمایند و با استناد به آن به طور مستمر کیفیت تجربه مشتری را اندازه گیری می نمایند تا در صورت تراکم شبکه و یا احتمال بروز قطعی، هشدارهای لازم به تیم های نگهداری و پشتیبانی داده شود . شرکت های ارایه دهنده خدمات ارتباطی قبلا از داده های گذشته جهت مدیریت شبکه استفاده می کردند ولی با تجزیه و تحلیل های کلان داده و پردازش بلادرنگ داده شبکه می توان از یک طرف وضعیت شبکه را بطور مستمر مانیتور کرد و ازطرف دیگر فعالیت های شبکه را مدل سازی و نیازهای آینده را مدیریت کرد . در نتیجه ، مهندسين شبکه می توانند یک دید کامل از رویدادهایی که در شبکه اتفاق می افتد را داشته باشند و بطور فعالانه به خرابی ها و قطعی های احتمالی شبکه پاسخ دهند تا در عمل میلیون ها دلار صرفه جویی گردد . مثلا ، شرکت های ارایه دهنده خدمات ارتباطی می توانند تاثیرات بالقوه (احتمالی ) را به صورت بلادرنگ برای یک سایت سلولی غیرفعال شده بر اساس تعداد مشترکین و ظرفیت سایت های همجوار ، مدل سازی نمایند. همچنین ، بر اساس داده جمع آوری شده بلادرنگ از برج های سلولی ، مهندسين می توانند هر گونه افت عملکرد سرویس در یک مکان بخصوص را مانیتور نمایند و با اعزام تیم های فنی ، در صورت نیاز، به صورت پیشگیرانه نسبت به حل مشکل اقدام نمایند .

#### ۴. نشت درآمد و تضمین درآمد :

برآورد شده است شرکت های ارایه دهنده سرویس های ارتباطی حدود ۸ / ۲ درصد از درآمد خود را هر سال به دلیل تقلب و یا نشت درآمد از دست می دهند که تقریبا معادل ۴۰ میلیارد دلار درآمد از تمامی صنعت تلکام است . این بدان معنی است که شرکت های تلکام می توانند بدون فروش محصولات و یا سرویس های اضافه، ۴۰ میلیارد دلار به درآمد خود اضافه نمایند ، مشروط به این که بتوانند با بکارگیری فناوری های نوین با نشت درآمد مقابله نمایند. شرکت های تلکام می توانند با استفاده از راه حل های کلان داده ، نقاط مستعدی که دارای پتانسیل بالقوه نشت درآمد می باشند را از طریق شبکه سیستم های در ارتباط مستقیم با مشتری تشخیص و قبل از ارسال اطلاعات به سیستم صورت حساب ، داده مربوطه را تصحیح نمایند . راه حل های مبتنی بر هدوپ می توانند به ارایه دهندگان سرویس های ارتباطی کمک کنند تا هم داده

ساخت یافته و هم داده غیرساخت یافته را پردازش و تحلیل نمایند و با بررسی اطلاعات در سال های گذشته در مقابل چندین ماه گذشته ، درک بهتری از رفتار مشتریان داشته باشند. در حال حاضر با استفاده از راه حل های مبتنی بر هدوپ و بکارگیری روش هایی نظیر بازرسی عمیق بسته یا DPI (برگرفته شده از Deep Packet Inspection ) به منظور تشخیص تقلب و نشت درآمد و همچنین شناسایی فرصت های جدید درآمدی ، این کار مقرون به صرفه شده است . فرآیند DPI ، حجم بالایی از داده را تولید می کند و میلیون ها رکورد در هرثانیه پردازش می شود ، کاری که قبلا بدون استفاده از راه حل نوین کلان داده ها نظیر هدوپ امکان جمع آوری و تحلیل آنها وجود نداشت .

#### ۵. مدیریت اطلاعات و امنیت سایبری :

به موازات گسترش دستگاه ها ، امنیت سایبری برای شرکت های تلکام در کانون توجه جدی قرار گرفته است . ایمن بودن شبکه و سیستم های مرتبط ارایه دهندگان سرویس های ارتباطی در مقابل حملات کدهای مخرب به یکی از جنبه های حیاتی کسب و کار آنها تبدیل شده است . قابلیت های مرسوم تشخیص رویدادها ، قادر به جمع آوری و تحلیل تمامی منابع داده ضروری برای شناسایی و پاسخ گویی به تهدیدات پیشرفته با توجه به پیچیدگی و هزینه های مرتبط نمی باشند. کارشناسان حرفه ای به منظور کاهش ریسک ، تشخیص رویدادها و پاسخ به رخنه ها و آسیب ها ، لازم است به حجم بالایی داده ( شامل لاگها ، رویدادها ، پکت ها ، جریان داده ، داده های مهم ، داده پیکربندی و ... ) دستیابی و آنها را به صورت بلادرنگ تحلیل نمایند تا ضمن کاهش سطح خطرات ، قدرت تشخیص به موقع وقایع امنیتی و رخنه های امنیتی و پاسخ مناسب به آنها را افزایش دهند. ارایه دهندگان سرویس های ارتباطی بطور فزاینده ای به سراغ راه حل های کلان داده رفته اند تا با بکارگیری آنها اقدام به جمع آوری و تحلیل لاگ های داده ، یافتن موارد غیرطبیعی (هشدارنسبت به مواردغیرطبیعی) و ایجاد یک رویداد برای یک تحلیل امنیتی نمایند. هاب های داده مبتنی بر هدوپ قادر به ارایه پلت فرم های مقرون به صرفه برای ذخیره سازی و قابلیت های تحلیل پیشرفته به منظور حمایت تحلیل های عمیق packet ، تحلیل های رفتاری ، پروفایلیگ و مدل سازی تهدیدات هستند.

#### ۶. تجزیه و تحلیل به عنوان سرویس (DAaaS) :

با ترکیب اطلاعات مکان مشتری با جمعیت شناسی و اولویت ها ، شرکت های ارایه دهنده سرویس های ارتباطی به دنبال ارایه تجزیه و تحلیل به عنوان یک سرویس (DAaaS) به سایر کسب و کارها در حوزه های مختلف دیگری نظیر خرده فروشی ها ، سرویس های مالی ، تبلیغات ، سلامت و درمان ، سرویس های عمومی و سایر کسب و کارهای در ارتباط مستقیم به مشتری می باشند . مجموعه ای گسترده از کاربردها و موارد کاری برای تحلیل های داده محور وجود دارد . شناخت الگوهای ترافیک شهری و گلوگاه ها ، کمک به شرکت های تدارکاتی جهت تنظیم دقیق فرآیندهای عرضه محصولات و کمک به شرکت های تبلیغاتی برای ارایه کمپین های هدفمند و تبلیغات برای بخش های خاص و کوچک تری از مشتریان ، صرفا نمونه اندکی در این رابطه می باشند. برخی از شرکت های پیشرو ارایه دهنده سرویس های

ارتباطی نظیر Sprint، Verizon و Telefonica در حال سرمایه گذاری بر روی این فرصت ها می باشند و موجودیت های کسب و کاری را ایجاد کرده اند که تمرکز آنها بر ارایه و عرضه سرویس های تحلیلی و کسب درآمد از طریق داده برای سایر کسب و کارها است.

با این که بازار فرصت های تحلیل داده همچنان نوپا و در ابتدای کار است ، کشش قابل توجه و اشتیاق فراوانی در این فضای کسب و کار وجود دارد و شرکت های ارایه دهنده خدمات ارتباطی می توانند با شتاب مناسب به سمت افزایش سهم خود از این نوع سرویس های تحلیلی در آینده باشند . علی رغم موارد فوق ، شرکت های تلکام می بایست به مسائل حریم شخصی و اطلاعات و مسائل مرتبط با آنها نیز توجه کنند و با تجمع و ناشناس کردن این اطلاعات ( داده مشترک) مطمئن شوند اطلاعات محرمانه در خصوص مشتریان خود را فاش نکرده اند. گرچه ، علی رغم فرصت های طلایی اشاره شده ، حفظ حریم خصوصی همچنان یک نگرانی جدی است و می بایست شرکت های ارایه دهنده سرویس های ارتباطی تجمع و گمنام سازی داده مشتریان را با دقت و وسواس دنبال نمایند تا این اطمینان حاصل گردد که اطلاعات محرمانه مشتریان افشا نمی شود . به هر حال ، حفظ حریم خصوصی مشتریان و محرمانگی یک نگرانی جدی است که اگر به درستی با آن برخورد شود ، شرکت های ارایه دهنده سرویس های ارتباطی دارای این فرصت طلایی خواهند بود که از طریق فروش بیش مشتریان به سایر کسب و کار های مرتبط بدون به خطر انداختن حریم خصوصی و حقوق مشترکین منابع جدید درآمدی برای خود ایجاد نمایند .

## - نقش کلان داده ها در بازاریابی

بسیاری از صنایع در تلاش برای تشخیص سریع تر اهداف مشتری ها، با حقیقتی نگران کننده روبرو شده اند. تکنیک های مورد نظر برای پیش بینی خریدها و جذب مشتریان، تنها برای بازه های کوتاه مدت نتیجه بخش است. در این بازی، رقابت بازاریاب ها برای شناسایی حرکت بعدی مشتریان به توازن رسیده و تلاش ها بر کسب سود حاشیه ای متمرکز شده اند. نتیجه، رقابتی بی برنده در بازه ای زمانی کوتاه مدت است. همه در بازه های زمانی میان مدت و بلندمدت دوشادوش هم در حرکت به سوی برابری پیش رفته و هیچ یک از مدعیان مزیت رقابتی پایداری برای پیروزی در اختیار ندارند.

البته معنی گفته های بالا، توقف تلاش شرکت ها برای پیش بینی بهینه تر رفتار مشتریان نیست؛ اما زمانی که همه ی رقبا مشغول همین تلاش هستند و در شرایطی که ماندن در بازی هم پاداش های خودش را برای صنایع مختلف دارد، حداکثر نتیجه ای که از این تلاش به دست می آید، کسب درآمد بالاتر از حد میانگین خواهد بود. در بسیاری صنایع از جمله گردشگری، بیمه، ارتباطات، موسیقی و حتی خودروسازی، تکنیک های بازاریابی به حدی پیشرفت کرده که رقبا به سطحی برابر در قدرت جذب مشتری رسیده اند. مهم ترین و تنها مولفه ی باقیمانده برای جذب مشتری بیشتر، قدرت پیش بینی رفتار مصرف کننده های نهایی است.

برای دستیابی به این مزیت رقابتی، برنامه‌های بازاریابی، به تحلیل اطلاعات کلان داده‌ها برای پاسخ به سوالات استراتژیک درباره‌ی چگونگی جذب بیشتر مشتریان و ایجاد حس وفاداری و برقراری ارتباط پایدار با آن‌ها، نیاز دارند؛ ارتباطی که بتواند با پایداری در بلندمدت، نیازهای مشتریان را در مقابل رقبا، مرتفع کند.

سوال‌های استراتژیک یادشده، به موارد مهمی همچون کشف روش‌های ایجاد حس وفاداری در مشتریان کمک شایانی می‌کنند. در شرایط فعلی، قیمت‌گذاری مناسب برای جذب بیشتر مشتری، کارساز نیست، بلکه باید راهی برای ایجاد ارزش بیشتر در محصولات و خدمات پیدا کرد که مشتریان قدر آن را بدانند. یافتن روش‌هایی برای عدم رغبت مشتریان به امتحان کردن یا انتخاب محصول و خدمات رقیب، استراتژی‌ای بسیار مهم‌تر از تبدیل مشتریان رقیب به مصرف‌کنندگان کالاها یا خدمات خود است.

پاسخ به این سوالات اساسی و یافتن راه‌حل‌های مطلوب، نیازمند استفاده‌ی متفاوت از کلان داده‌ها است. لازم است به جای یافتن راه‌حلی برای هدف‌گیری مشتریان بالقوه، به دنبال استفاده از کلان داده‌ها جهت کسب صفات ارزشمند در محصولات و خدمات بود. این صفات ارزشمند، چیزهایی هستند که مشتریان برای انتخاب گزینه‌ی ایده‌آل به آن توجه می‌کنند. سوال اساسی و متفاوت این است: کلان داده‌ها چگونه می‌توانند به ما - و نه مشتریان - کمک کنند؟

با کمک کلان داده‌ها می‌توان اطلاعاتی جدید را برای ایجاد محصولات و خدماتی کاملاً تازه کشف کرد. مثال ساده در همین رابطه، استفاده از کلان داده‌ها برای کمک به تسریع روند یافتن نتیجه‌ی مورد نظر کاربران به هنگام استفاده از موتورهای جست‌وجو باشد. روشی که در حال حاضر شرکت‌هایی مانند آمازون و نت‌فلیکس برای کمک به کاربران‌شان جهت یافتن سریع‌تر گزینه‌های مورد نظرشان، از آن بهره می‌برند. در مثالی دیگر می‌توان به تجمع منابع اطلاعاتی‌ای که شرکت‌ها از نتیجه‌های جست‌وجوی سایر مشتریان و کاربران‌شان به‌دست می‌آورند اشاره کرد. اطلاعاتی که با کمک آن‌ها می‌توان با تشخیص روند جست‌وجوی سایر مشتریان و مقایسه‌ی آن با مشتری فعلی، نتایج بهتری را در اختیار او گذاشت.

## - جمع بندی :

با ورود به عصر اطلاعات و ارتباطات و آغاز استفاده از داده‌ها و اطلاعات به عنوان سرمایه‌های اصلی در حرکت علمی، اقتصادی، اجتماعی و فرهنگی جوامع، سازمان‌ها و شرکت‌های مختلف و توسعه مشارکت افراد در جهان اینترنت و ارتباطات شبکه‌ای در دنیا، دغدغه‌ای بروز پیدا کرد که از جنس همین داده‌هایی بود که همه روز و با سرعت وحشتناک در دنیا و در عرصه‌های مختلفی که فناوری اطلاعات ورود پیدا کرده بود، تولید می‌شود و آن اینکه چگونه این حجم بزرگ و متنوع داده‌ها و اطلاعات را با توجه به ساختارهایی که در فضای فناوری اطلاعات وجود دارد، می‌توان مدیریت، کنترل و پردازش کرد و از آن در جهت بهبود ساختارها و سودآوری بیشتر بهره جست؟

از سال ۲۰۱۲ به بعد در هر روز هزار پتا بایت (Pebibyte ۱۰۰۰) داده تولید می شود که به دنبال خود مستلزم ذخیره سازی، تحلیل، جستجوها، تمیزکاری داده ها، اشتراک ها و ... در داده هاست که باید در حوزه مختلف انجام شود.

این موضوع باعث شده است که پژوهشگران و دانشمندان به دنبال ایجاد ساختارها، متدلوژی ها، روش ها و رویکردهای جدیدی برای مدیریت، کنترل و پردازش این حجم از داده ها باشند که این تلاش ها در ذیل سایه «داده های بزرگ» مطرح شده است.

داده های بزرگ نحوه کار سازمان ها و افراد را تحت تاثیر قرار می دهد. داده های بزرگ فرهنگی را در سازمان ها ایجاد می کند که از طریق آن کسب و کارها و مدیران فناوری اطلاعات را به سمت استفاده از تمامی ارزش های پنهان در داده ها سوق می دهد. ادراک این ارزش ها به همه کارکنان سازمان ها این امکان را می دهد که با بیش وسیع تری تصمیم گیری کنند، نزدیکی بیشتری با مشتریان داشته باشند، فعالیت های خود را بهینه کنند، با تهدیدات مقابله کنند و در نهایت سرمایه های خود را بر روی منبع جدیدی از سود سرشار پنهان در داده ها متمرکز سازند. سازمان ها برای رسیدن به این مرحله نیازمند معماری جدید، ابزارهای نو و فعالیت ها و تلاش های مستمری هستند تا بتوانند از مزیت های چهارچوب های مبتنی بر داده های بزرگ بهره مند گردند.

با توجه به اینکه در این پژوهش، هدف تجزیه و تحلیل کاربرد Big Data در فرآیندهای بازاریابی می باشد، بدین منظور ابتدا متغیرهایی که بر فرآیند بازاریابی تاثیرگذار می باشد را با مطالعه ی تحقیقاتی نظیر مغنی و همکاران (۱۳۹۴)، قورچیان و همکاران (۱۳۹۳)، ذوالفقاری (۱۳۹۲)، سهراب نژاد (۱۳۹۱) بررسی کرده، سپس ارتباط بین Big Data و این متغیرها را به عنوان فرضیه آورده شد.

#### آزمون فرضیات

- فرضیه ۱: بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد.
- فرضیه ۲: بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.
- فرضیه ۳: بین کاربرد Big Data و تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی رابطه معناداری وجود دارد.
- فرضیه ۴: بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.
- فرضیه ۵: بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

## پیشینه‌ی تحقیق

### پیشینه‌ی داخلی

۱. عنایتی در سال ۱۳۹۳، در تحقیقی به بررسی کاربرد رایانش ابری در داده‌های حجیم پرداختند. داده‌های حجیم یک اصطلاح برای مجموعه داده‌های خیلی بزرگ است که از نظر ساختار، پیچیدگی و منابع تولید بسیار متنوع هستند و ذخیره و آنالیز آنها کار پیچیده‌ای است. رایانش ابری یک تکنولوژی قدرتمند برای اجرای محاسبات پیچیده و سنگین است. رایانش ابری نیاز به استفاده از سخت‌افزارهای گران را حذف نموده و فضای محاسباتی و نرم‌افزار مورد نیاز را در اختیار کاربر قرار می‌دهد. رشد روزافزون حجم داده و ایجاد داده‌های حجیم از طریق رایانش ابری در سالهای اخیر در بسیاری از کاربردها دیده شده است. داده‌های حجیم چالش مهمی است که احتیاج به زیرساختی قوی برای اطمینان از انجام موفق پردازشها و آنالیزهای مورد نیاز دارد. موضوع حایز اهمیت این است که چگونه میتوان از زیرساخت رایانش ابری برای دسترسی، پردازش و آنالیز داده‌های حجیم استفاده نمود. در این مطالعه به تعریف، خصوصیات و دسته‌بندی داده‌های حجیم در چارچوب رایانش ابری پرداخته شده است و کاربرد زیرساخت رایانش ابری برای آنالیزهای داده‌های حجیم مورد بررسی قرار گرفته است. نتایج این تحقیق نشان‌دهنده‌ی این بود که در حال حاضر داده از نظر اندازه در حال بزرگ شدن است و این روند رو به رشد با افزایش تنوع داده تولید شده بیشتر میشود. سرعت تولید داده به دلیل استفاده زیاد از وسایل همراه و حسگرهای متصل به اینترنت در حال افزایش است. داده‌های تولید شده فرصتی مناسب برای همه صنایع و حرفه‌ها ایجاد میکنند تا با آنالیز داده‌های حجیم به آگاهی بهتر نسبت به کسب و کار خود دست یابند. امروزه سرویسهای ابری برای ذخیره، پردازش و آنالیز داده‌های محیطی مناسب هستند. این سرویس‌ها چهره تکنولوژی‌های ارتباطی را تغییر داده‌اند.

۲. رضاپور نیاری در سال ۱۳۹۶، در تحقیقی به بررسی کاربرد کلان داده در اینترنت اشیا و تأثیر آن بر صنعت برودکست پرداخت. اینترنت اشیا، فناوری پیشرفته‌ای است که در آن، هر موجود (انسان و غیرانسان) امکان تبادل داده با شبکه‌ای از اشیا را دارد. موجودات در این فناوری دارای هویت مستقل بوده و امکان مدیریت و ساماندهی آنها از طریق کامپیوترها فراهم است. درون هر یک از اشیای معرفی شده سنسورهایی تعبیه شده است که به تولید و ارسال داده می‌پردازد. با توجه به روند روبه‌رشد دستگاه‌های هوشمند و اتصال این دستگاه‌ها



به شبکه‌ی اینترنت، حجم داده‌های تولید شده روزبه‌روز در حال افزایش است و لازم است از تکنولوژی‌های کلان داده برای تحلیل و استخراج دانش و بینش موجود در داده‌های مبادله شده استفاده کرد. بکارگیری دانش مذکور می‌تواند موجب افزایش کارایی و بهره‌وری در صنایع مختلف اعم از ساخت و تولید، سلامت، حمل‌ونقل، مصرف انرژی و برودکست شود. استفاده از این فناوری در صنعت برودکست می‌تواند علاوه بر تأثیر بر رونق کسب‌وکار و نحوه درآمدزایی از محتوا، از طریق خدماتی نظیر ارائه محتوا در هر مکان، زمان و در چند اسکرین، ارائه پیشنهاد در خصوص محتوای مدنظر مخاطب، اشتراک تجربیات و نظرات مخاطبان، مدیریت بسته‌های تبلیغاتی، ارائه خدمات ارزش افزوده، موجب افزایش رضایت مخاطبان، حفظ مخاطبان فعلی و جذب بیشتر مخاطبان می‌گردد.

۳. نوریان و حجازی در سال ۱۳۹۲، در تحقیقی به بررسی کاربرد کلان‌داده‌ها در نقد توسعه مبتنی بر حمل و نقل عمومی پرداختند. امروزه با پیشرفت فناوری اطلاعات و تکنولوژی دیجیتال، حجم بالایی از داده‌ها با عناوین کلان-داده و داده‌های بهنگام، در اختیار متخصصان علوم مختلف مهندسی و نیز علوم اجتماعی قرار گرفته است که از طریق آن امکان ارزیابی نظریه‌های موجود با قابلیت تعمیم‌پذیری فراهم شده است. یکی از این رویکردها در شهرسازی، توسعه مبتنی بر حمل و نقل عمومی است که بر اساس سفرسازی کاربری‌ها و ارقام موجود حاصل از مدل‌های حمل و نقل شهری پیشنهادهای برنامه‌ریزی کاربری زمین و تغییرات تراکم را مطرح می‌سازد. در این پژوهش از تفسیر کلان-داده‌ها برای اعتبارسنجی معیارهای تجربی مربوط به توسعه مبتنی بر حمل و نقل عمومی و برنامه‌ریزی حمل و نقل شهری در خصوص سفرسازی کاربری‌ها استفاده می‌شود. بدین منظور الگوی ترافیکی کاربری آموزش عالی با روش ارزیابی همبستگی تحلیل می‌گردد. با بکارگیری نگرش شهرسازی در پالایش کلان-داده‌های ترافیکی موجود و در دسترس برای دانشگاه ایالتی کالیفرنیا در شهر لس آنجلس، الگوی ترافیکی سفرهای تحصیلی استخراج می‌شود. نتایج ارزیابی الگوی سفرهای تحصیلی نشان می‌دهد، تولید سفر بر ارقام سفرسازی کاربری‌ها استوار نبوده، بلکه این الگو تابعی از ترافیک کل شهر است. این نتایج، طراحی و برنامه‌ریزی کاربری زمین را که بر اساس مدیریت تقاضا در مدل‌های حمل و نقل شهری و بخصوص توسعه شهری مبتنی بر حمل و نقل عمومی انجام می‌شود، با تردید جدی

مواجه می سازد. یافته های این پژوهش، مثالی از کاربرد کلان داده ها به عنوان یکی از دستاوردهای بروز فناوری اطلاعات در مطالعات شهرسازی است.

۴. خدیور و امیریان در سال ۱۳۹۵، در تحقیقی به بررسی ارزیابی بلوغ کلان داده در صنعت بیمه کشور پرداختند. با توجه به رشد روز افزون داده و نیاز به تحلیل و استفاده از این حجم بالای داده، به راهبردهای خاص مدیریتی نیاز هست، که در دنیای فناوری اطلاعات به عنوان کلان داده شناخته می شود. این موضوع، به این دلیل هر روز جذابیت و مقبولیت بیشتری پیدا میکند که با استفاده از تحلیل حجم های بیشتری از داده ها، می توان تحلیل های بهتر و پیشرفته تری را برای مقاصد مختلف، از جمله مقاصد تجاری، پزشکی و امنیتی، انجام داد و نتایج مناسبتری را دریافت کرد. برای ایجاد یک دید مناسب در خصوص کلان داده و اهمیت آن، جامعه ای را تصور کنید که در آن جمعیت بطور نمایی در حال افزایش است، اما خدمات و زیرساخت های عمومی آن نتواند پاسخگوی رشد جمعیت باشد و از عهده مدیریت آن برآید. چنین شرایطی در حوزه داده در حال وقوع است. بنابراین نیازمند توسعه زیرساخت های فنی برای مدیریت داده و رشد آن در بخش هایی نظیر جمع آوری، ذخیره سازی، جستجو، به اشتراک گذاری و تحلیل می باشیم. با توجه به اهمیت صنعت بیمه در اقتصاد کشور و مواجهه این صنعت با حجم زیادی از داده های ساختار یافته و ساختار نیافته در این پژوهش به ارزیابی بلوغ کلان داده در صنعت بیمه کشور پرداخته شده است و از آنجایی که تاکنون ارزیابی کلان داده در صنعت بیمه کشور انجام نشده است و وضعیت بلوغ کلان داده در این صنعت مشخص نمی باشد، با این پژوهش سعی بر آن شده است که وضعیت بلوغ کلان داده در کلیه شرکت های بیمه فعال در ایران مشخص گردد. روش تحقیق مورد استفاده توصیفی - پیمایشی می باشد و از پرسشنامه استاندارد برای ارزیابی بلوغ کلان داده استفاده شده است جامعه آماری متشکل از تمام شرکت های بیمه است، نمونه این پیمایش از بین افراد مطلع، خبرگان و مدیران حوزه فناوری اطلاعات کلیه شرکت های بیمه در کشور ایران گرفته شده است که تعداد آن ۵۱ عدد می باشد. با توجه به ارزیابی بدست آمده ۵۵ شرکت بیمه از نظر وضعیت بلوغ کلان داده، نوپا می باشند و ۵ شرکت در وضعیت پیش از پذیرش هستند و ۵ شرکت در مرحله پذیرش اولیه می باشند. بر اساس بررسی صورت گرفته، متناسب با وضعیت هر شرکت پیشنهاداتی داده شده است.

۵. محمودزاده و همکاران در سال ۱۳۹۶، در تحقیقی به بررسی تدوین استراتژی کلان داده در تحلیل شبکه های اجتماعی برای پیش بینی بحران پرداختند. امروزه به دلیل پیچیدگی های فضای سایبری و تاثیرگذاری آن بر بسیاری حوزه های دیگر از منظر امنیت ملی نیاز به سیستم پایش با رویکرد روندیابی فضای سایبری و نیز داشتن دید جامع نسبت به همه ی ابعاد آن شامل فرهنگی، اجتماعی، سیاسی، اقتصادی، نظامی امنیتی و علم و فناوری بیش از پیش ضروری است. با گسترش سیستم های پایگاهی و حجم بالای داده های ذخیره شده در این سیستم ها، نیاز به ابزاری است تا بتوان داده های ذخیره شده را پردازش کرد و اطلاعات حاصل از این پردازش را در اختیار کاربران قرار داد. با توجه به گسترش روزافزون ابزارهای ارتباطی در حوزه ی فناوری اطلاعات و ارتباطات و لزوم ایجاد آگاهی اطلاعاتی بر این حوزه با رویکرد جلوگیری از غافل گیری در مواجهه با بحران و از سوی دیگر حجم، سرعت و تنوع داده های شبکه های اجتماعی، لزوم تدوین استراتژی فناوری کلان داده در تحلیل شبکه های اجتماعی امری بسیار ضروری است. گفتنی است این مفهوم به جهت تازگی با روش های تحلیلی سنتی به هیچ وجه قابل اجرا نیست و باید در مواجهه با این حوزه از فناوری های نوین بهره برد. از آنجایی که یکی از اصلی ترین مباحث در حوزه ی کلان داده روش های داده کاوی برای پالایش داده و دست یابی به ارزش افزوده در حوزه ی داده های عظیم است، در پژوهش حاضر سعی شده است تا فناوری های مورد نیاز شناسایی شده و ضمن تحلیل و اولویت بندی آن به تدوین استراتژی فناوری کلان داده در تحلیل شبکه های اجتماعی برای پیش بینی وقوع بحران پرداخته شود. با توجه به موضوع و هدف پژوهش، نوع تحقیق کاربردی - توسعه ای ۱ است. روش تحقیق کمی است و جمع آوری اطلاعات از طریق پانل خبرگان و استفاده از پرسشنامه ی بسته بوده است و جامعه ی آماری نیز تعداد ۲۰ نفر از نخبگان عرصه ی فناوری اطلاعات و ارتباطات بوده است و روش تحلیل بهره گیری از طیف لیکرت برای ارزش گذاری اولویت های استراتژی کلان داده بوده است .

۶. هلیلی و همکاران در سال ۱۳۹۵، در تحقیقی به بررسی لزوم بکارگیری فناوری کلان داده در سامانه های C4I و بررسی چالش های آن پرداختند. کلان داده برای اشاره به داده هایی با حجم بسیار بالا، سرعت تولید و انتقال زیاد و متنوع بکار می رود؛ که با تجهیزات رایج قابل کنترل و مدیریت نیستند. این فناوری کاربردهای گستردهای در حوزه های هواشناسی، پزشکی، تجارت، مهندسی اجتماعی، صنایع و سازمان های دفاعی - امنیتی دارد. سامانه C4I به عنوان هسته مرکزی در فرماندهی و کنترل نظامی، مدیریت بحران و پدافند غیرعامل در شرایط اضطراری بکار می رود. در این سامانه، با توجه به حجم بالا و اهمیت سرعت پردازش و تبادل اطلاعات، استفاده از فناوری کلان داده به منظور تسریع در چرخه عملیاتی، امری ضروری به نظر می رسد. در این مقاله با مروری بر سامانه های نوین C4I، به معرفی ویژگی ها، کاربردها، ابزار و فناوری های کلان داده

پرداخته شده و چالش های مربوط به تعامل این دو حوزه بررسی شده است. شناخت این دو فناوری و چالش های مربوط به تلفیق آنها، زمینه ساز بر نامه ریزی مناسب برای بهره گیری از فرصت های بکارگیری کلان داده در سامانه ها C4I در کشور خواهد بود.

۷. وکیلی و کارگر در سال ۱۳۹۳، در تحقیقی به بررسی بهبود توازن بار در مدل نگاشت کاهش جهت پردازش کلان داده در خوشه های ناهمگن پرداختند. در سالهای اخیر، مدل های برنامه نویسی مختلفی برای پردازش کلان داده ارائه شده اند. در این میان، مدل نگاشت کاهش - ۲ توانسته است محبوبیت زیادی، به ویژه در محیط های رایانش ابری بدست آورد. نگاشت کاهش علاوه بر مزایا و ویژگی های منحصر به فردی که دارد، طبیعتاً دارای - چالش های مختلفی نیز هست. مسئله توازن بار یکی از چالش های مهم در مدل های برنامه نویسی توزیع شده از جمله نگاشت کاهش است، و در - صورتی که سازوکار مناسبی برای آن اتخاذ نشود می تواند باعث افزایش زمان اجرای برنامه ها و در نتیجه کاهش کارایی شود. از جمله معیارهایی که در توازن بار نقش کلیدی دارند، محلیت داده ها و انحراف داده ها هستند. ضمن اینکه برای برقراری توازن بار، باید به همگن یا ناهمگن بودن خوشه های محاسباتی نیز توجه داشت. در این تحقیق، روشی بر اساس کنترل ظرفیت ارائه می شود که توازن بار را در مدل نگاشت کاهش برای - خوشه های ناهمگن بهبود میبخشد. بر اساس ارزیابیها و آزمایش های انجام شده، روش پیشنهادی با افزایش محلیت داده ها و در نتیجه بهبود توازن بار، زمان اجرا را به طور میانگین حدود ۲۰٪ بهبود بخشیده است.

۸. آذریان در سال ۱۳۹۶، در تحقیقی به بررسی چالش امنیت کلان داده در فناوری ابر پرداختند. در سال های اخیر شاهد انفجاری از داده ها از طریق رسانه های اجتماعی، وب سایت ها، اینترنت اشیا، موبایل ها، ایمیل ها و هزاران فناوری نوین دیگر بوده ایم که بی گمان ضبط، نگهداری و تحلیل این داده های کلان و پرشتاب می تواند دارایی های ارزشمندی را تولید نماید. چالش چگونگی ذخیره سازی این کلان داده ها و تبدیل به نتایج سودمند و از سویی دیگر تامین هزینه زیرساخت لازم و نیروی خبره ی کار با این وسعت از داده ها، صاحبان کلان داده را به سمت استفاده از فناوری ابر سوق داده است. فناوری ابر، با فراهم آوری زیرساخت لازم، نیروی متخصص و سرویس های مورد نیاز برای نگهداری و پردازش کلان داده و سپس به اشتراک گذاری و یا اجاره دادن این امکانات به صاحبان کلان داده، آن ها را در رفع چالش های موجود یاری می کند. مانع بزرگ همکاری این دو فناوری قدرتمند و پیشرو، بحث پراهمیت امنیت و محافظت از داده های حساس و یا شخصی افراد، سازمان ها و دولت ها در هنگام برون سپاری این خدمات به تامین کنندگان ابر است. در این مقاله سعی

شده با تمرکز بر بیان چالش های امنیتی موجود، راهکارهای کاهش این مخاطرات و تهدیدات امنیتی را بررسی شود.

۹. علی دوست در سال ۱۳۹۴، در تحقیقی به بررسی اهمیت کلان داده و کاربردهای آن در حوزه های مختلف پرداختند. با توجه به رشد روزافزون داده ها نیاز به بهره برداری و تحلیل از این حجم بالای داده ها؛ نیاز به راهبردهای خاص مدیریتی دارد که در دنیای فناوری اطلاعات به کلان داده معروف است. داده های بزرگ یا کلان داده ترجمه اصطلاح Big Data می باشد که معمولاً به مجموعه ای از داده ها اطلاق می شود که اندازه آنها فراتر از حدی است که با نرم افزارهای معمول بتوان آنها را در یک زمان معقول مدیریت و پردازش کرد. در این مقاله به ویژگی های کلان داده فناوری های نوین پردازشی که امروزه برای مدیریت و تحلیل کلان داده مورد استفاده قرار می گیرد و سپس به کاربردهای کلان داده و اهمیت آن نیز اشاره شده است.

### پیشینه ی خارجی

۱- لابریدنس و جگادیش در سال ۲۰۱۲ در تحقیقی به بررسی چالش ها و فرصت ها با داده های بزرگ پرداختند. در حال حاضر وعده تصمیم گیری مبتنی بر داده ها به طور گسترده به رسمیت شناخته شده است و شور و شوق رو به رشدی برای مفهوم "بزرگ داده ها"، از جمله اعلام اخیر از سوی کاخ سفید در مورد طرح های جدید مالی در میان آژانس های مختلف وجود دارد، که هدف پژوهش برای بزرگ داده ها می باشد. به عنوان مثال، برآورد شده است که گوگل به تنهایی ۵۴ میلیارد دلار برای اقتصاد ایالات متحده در سال ۲۰۰۹ مساعدت کرده است. به نظر می رسد هیچ توافقی روشنی در مورد ماهیت بزرگ داده ها وجود ندارد. در حقیقت، بیانیه های بحث برانگیز فراوانی در مورد داده های بزرگ، مانند "اندازه تنها چیزی است که مهم است" وجود دارند. در این طرح سعی خواهیم شد تا اختلافات را کشف و اسطوره های اطراف بزرگ داده را کم ارزش شود.

۲- مک کینزی و همکاران در سال ۲۰۱۱، در تحقیقی به بررسی داده های بزرگ: مرز بعدی برای نوآوری، رقابت و بهره وری پرداختند. اطلاعات در حال انفجار در جهان هستند و تحلیل مجموعه داده های بزرگ - به اصطلاح بزرگ داده - تبدیل به یک مبنای کلیدی رقابت خواهد شد، موجب موج جدیدی از رشد بهره وری، نوآوری و مازاد مصرف کننده بر اساس تحقیقات MGI و دفتر فناوری کسب و کارمکنزی را در می نوردد. مدیران در هر بخش - نه فقط چندین مدیر اطلاعاتی - باید با مفاهیم داده

های بزرگ روبرو شوند. حجم در حال افزایش و جزئیات اطلاعات گرفته شده توسط شرکت ها، رشد رسانه های چندرسانه ای، رسانه های اجتماعی و اینترنت چیزها، رشد چشمگیر داده ها را در آینده قابل پیش بینی تحریک خواهند کرد.

۳- بیگولی و هوری در سال ۲۰۱۲، در تحقیقی به بررسی معماری نرم افزار (WICSA) و کنفرانس اروپایی معماری نرم افزار (ECSA) پرداختند. پدیده داده های بزرگ به تمرکز جمع آوری و پردازش مجموعه داده های بسیار بزرگ و سیستم های مرتبط و الگوریتم های مورد استفاده برای تجزیه و تحلیل این مجموعه داده های عظیم اشاره دارد. معماری برای داده های بزرگ معمولاً در میان ماشین آلات و خوشه های مختلف قرار دارد و معمولاً شامل چندین ریزسیستم با اهداف خاص می باشند همراه با فرآیند کشف دانش، جنبش داده های بزرگ، فرصت های منحصر به فردی را برای سازمان ها به ارمغان می آورد (با توجه به بینش های جدید، بهینه سازی کسب و کار، و غیره). با این حال، با توجه به مشکل تجزیه و تحلیل چنین مجموعه داده های بزرگ، داده های بزرگ سیستم های مهندسی منحصر به فرد و چالش های معماری ارائه می دهند. در این مقاله، ما سه اصول طراحی سیستم را ارائه می دهیم که می توانند سازمان ها را در فرایندهای تحلیلی و جمع آوری داده های مؤثر، سازماندهی سیستم و شیوه های انتشار داده ها، مطلع کنند. اصول ارائه شده از تجربیات تحقیق و توسعه خود ما با مشکلات داده های بزرگ از ادارات مختلف فدرال حاصل می شود و ما هر اصل را براساس تجربه و توصیه های خود نشان می دهیم.

۴- ارولز و همکاران در سال ۲۰۱۶، در تحقیقی به تجزیه و تحلیل مصرف کننده کلان داده و تحول بازاریابی پرداختند. تجزیه و تحلیل مصرف کننده در مرکز زلزله انقلاب داده های بزرگ قرار دارد. فن آوری به جذب اطلاعات غنی و فراوان در مورد پدیده مصرف کننده در زمان واقعی کمک می کند. بنابراین، حجم بی سابقه، سرعت و انواع داده های اولیه، کلان داده، از تک تک مصرف کنندگان در دسترس هستند. برای درک بهتر اثر کلان داده در فعالیت های مختلف بازاریابی، شرکت ها قادر به بهره مندی بیشتر از مزایای خود هستند، چارچوب مفهومی که بر مبنای تئوری مبتنی بر منابع استوار است، پیشنهاد می شود. سه منبع سرمایه فیزیکی، انسانی و سازمانی عبارتند از: (۱) فرایند جمع آوری و ذخیره سازی شواهد فعالیت مصرف کننده به عنوان اطلاعات بزرگ، (۲) فرایند استخراج بینش مصرف کننده از اطلاعات بزرگ و (۳) فرآیند استفاده از بینش مصرف کننده برای افزایش قابلیت های پویا /

سازگاری. علاوه بر این، منابع مورد نیاز منحصر به فرد برای شرکت ها برای استفاده از داده های بزرگ مورد بحث قرار می گیرد.

۵- بوید و کراوفورد در سال ۲۰۱۲، در تحقیقی به بررسی سوالات انتقادی برای کلان داده ها پرداختند. انگیزش برای یک پدیده فرهنگی، تکنولوژیکی و علمی: دوران داده های بزرگ آغاز شده است. متخصصان کامپیوتر، فیزیکدانان، اقتصاددانان، ریاضیدانان، دانشمندان علوم سیاسی، متخصصان علوم زیستی، جامعه شناسان و سایر محققان برای دستیابی به مقادیر عظیمی از اطلاعات تولید شده در باره و توسط مردم و چیزها، و تعاملات آنها، به دنبال تقاضا آن هستند. گروه های گوناگون در مورد مزایای بالقوه و هزینه های تجزیه و تحلیل توالی های ژنتیکی، تعاملات رسانه های اجتماعی، پرونده های سلامت، لگاریتم های تلفن، سوابق دولتی و سایر اثرات دیجیتالی به جا مانده از افراد بحث می کنند. سوالات مهمی شکل می گیرند. آیا داده های تحقیق در مقیاس بزرگ به ما در ایجاد ابزارها، خدمات و کالاهای عمومی بهتر کمک می کند؟ یا آن را در موج جدیدی از تهاجمات حریم خصوصی و بازاریابی هجوم کننده راهنمایی می کند؟ آیا تجزیه و تحلیل داده ها به ما در درک جوامع آنلاین و جنبش های سیاسی کمک می کند؟ یا آیا برای ردیابی معترضین و سرکوب سخنرانی استفاده می شود؟ آیا این نحوه مطالعه ما در ارتباط با فرهنگ و جوامعی انسانی تغییر می دهد و یا پالت گزینه های تحقیق را محدود می کند و آیا معنای "تحقیق" را تغییر می دهد؟ با توجه به ظهور کلان داده ها به عنوان یک پدیده اجتماعی-فنی، ما استدلال می کنیم که لازم است به طور منتقدانه به فرضیه ها و سوگیری های آن بپردازیم. در این مقاله ما شش تحریک کننده ارائه دادیم که مکالمات در مورد مسائل اطلاعات بزرگ خودنمایی کنند. یک پدیده فرهنگی، تکنولوژیکی و علمی که متکی بر تکنولوژی، تجزیه و تحلیل و اسطوره شناسی است که علم معانی بیان تخیلی و غیرتخیلی را به وجود می آورد.

جدول روند تحقیق

ردیف	اسم نویسنده	سال	هدف	نتیجه
۱	عنایتی	۱۳۹۳	کاربرد رایانش ابری در داده‌های حجیم	در حال حاضر داده از نظر اندازه در حال بزرگ شدن است و این روند رو به رشد با افزایش تنوع داده تولید شده بیشتر میشود. سرعت تولید داده به دلیل استفاده زیاد از وسایل همراه و حسگرهای متصل به اینترنت در حال افزایش است.
۲	رضاپور نیاری	۱۳۹۶	کاربرد کلان داده در اینترنت اشیا و تأثیر آن بر صنعت برودکست	استفاده از این فناوری در صنعت برودکست می‌تواند علاوه بر تأثیر بر رونق کسب-وکار و نحوه درآمدزایی از محتوا، از طریق خدماتی نظیر ارائه محتوا در هر مکان، زمان و در چند اسکرین، ارائه پیشنهاد در خصوص محتوای مدنظر مخاطب، اشتراک تجربیات و نظرات مخاطبان، مدیریت بسته‌های تبلیغاتی، ارائه خدمات ارزش افزوده، موجب افزایش رضایت مخاطبان، حفظ مخاطبان فعلی و جذب بیشتر مخاطبان می-



گردد				
۳	نوریان و حجازی	۱۳۹۲	کاربرد کلان داده ها در نقد توسعه مبتنی بر حمل و نقل عمومی	این نتایج، طراحی و برنامه ریزی کاربری زمین را که بر اساس مدیریت تقاضا در مدل های حمل و نقل شهری و بخصوص توسعه شهری مبتنی بر حمل و نقل عمومی انجام می شود، با تردید جدی مواجه می سازد. یافته های این پژوهش، مثالی از کاربرد کلان-داده ها به عنوان یکی از دستاوردهای بروز فناوری اطلاعات در مطالعات شهرسازی است .
۴	خدایور و امیریان	۱۳۹۵	بررسی ارزیابی بلوغ کلان داده در صنعت بیمه کشور	بر اساس بررسی صورت گرفته ، متناسب با وضعیت هر شرکت پیشنهاداتی داده شده است.
۵	محمودزاده و همکاران	۱۳۹۶	بررسی تدوین استراتژی کلان داده در تحلیل شبکه های اجتماعی برای پیش بینی بحران	فناوری های مورد نیاز شناسایی شده و ضمن تحلیل و اولویت بندی آن به تدوین استراتژی فناوری کلان داده در تحلیل شبکه های اجتماعی

				برای پیش بینی وقوع بحران پرداخته شد
۶	هلیلی و همکاران	۱۳۹۵	بررسی لزوم بکارگیری فناوری کلان داده در سامانه های C4I و بررسی چالش های آن	با مروری بر سامانه های نوین C4I، به معرفی ویژگی ها، کاربردها، ابزار و فناوری های کلان داده پرداخته شد و چالش های مربوط به تعامل این دو حوزه بررسی شد. شناخت این دو فناوری و چالش های مربوط به تلفیق آنها، زمینه ساز بر نامه ریزی مناسب برای بهره گیری از فرصت های بکارگیری کلان داده در سامانه ها C4I در کشور خواهد بود.
۷	وکیلی و کارگر	۱۳۹۳	بررسی بهبود توازن بار در مدل نگاشت کاهش جهت پردازش کلان داده در خوشه های ناهمگن	بر اساس ارزیابیها و آزمایش های انجام شده، روش پیشنهادی با افزایش محلیت داده ها و در نتیجه بهبود توازن بار، زمان اجرا را به طور میانگین حدود ۲۰٪ بهبود بخشیده است.
۸	آذریان	۱۳۹۶	به بررسی چالش امنیت کلان داده در فناوری ابر	با تمرکز بر بیان چالش های امنیتی موجود، راهکارهای کاهش این مخاطرات و تهدیدات امنیتی را بررسی شد.

۹	علی دوست	۱۳۹۴	بررسی اهمیت کلان داده و کاربردهای آن در حوزه‌های مختلف	به ویژگی های کلان داده فناوری های نوین پردازشی که امروزه برای مدیریت و تحلیل کلان داده مورد استفاده قرار می گیرد و سپس به کاربردهای کلان داده و اهمیت آن نیز اشاره شد
۱۰	لابریدنس و جگادیش	۲۰۱۲	بررسی چالش ها و فرصت ها با داده های بزرگ	به نظر می رسد هیچ توافقی روشنی در مورد ماهیت بزرگ داده ها وجود ندارد. در حقیقت، بیانیه های بحث برانگیز فراوانی در مورد داده های بزرگ، مانند "اندازه تنها چیزی است که مهم است" وجود دارند
۱۱	مک کینزی و همکاران	۲۰۱۱	بررسی داده های بزرگ: مرز بعدی برای نوآوری، رقابت و بهره‌وری	مدیران در هر بخش - نه فقط چندین مدیر اطلاعاتی - باید با مفاهیم داده های بزرگ روبرو شوند. حجم در حال افزایش و جزئیات اطلاعات گرفته شده توسط شرکت ها، رشد رسانه های چندرسانه ای، رسانه های اجتماعی و اینترنت چیزها، رشد چشمگیر داده ها را در آینده قابل پیش بینی

				تحریک خواهند کرد
۱۲	بیگولی و هوری	۲۰۱۲	بررسی معماری نرم افزار (WICSA) و کنفرانس اروپایی معماری نرم افزار (ECSA)	با توجه به مشکل تجزیه و تحلیل چنین مجموعه داده های بزرگ، داده های بزرگ سیستم های مهندسی منحصر به فرد و چالش های معماری ارائه می دهند
۱۳	ارولز و همکاران	۲۰۱۶	تجزیه و تحلیل مصرف کننده کلان داده و تحول بازاریابی	سه منبع سرمایه فیزیکی، انسانی و سازمانی عبارتند از: (۱) فرایند جمع آوری و ذخیره سازی شواهد فعالیت مصرف کننده به عنوان اطلاعات بزرگ، (۲) فرایند استخراج بینش مصرف کننده از اطلاعات بزرگ و (۳) فرایند استفاده از بینش مصرف کننده برای افزایش قابلیت های پویا / سازگاری. علاوه بر این، منابع مورد نیاز منحصر به فرد برای شرکت ها برای استفاده از داده های بزرگ مورد بحث قرار می گیرد.
۱۴	بوید و کراوفورد	۲۰۱۲	بررسی سوالات انتقادی برای کلان-داده ها	شش تحریک کننده ارائه داده-شد که مکالمات در مورد مسائل اطلاعات بزرگ

<p>خودنمایی کنند. یک پدیده فرهنگی، تکنولوژیکی و علمی که متکی بر تکنولوژی، تجزیه و تحلیل و اسطوره شناسی است که علم معانی بیان تخیلی و غیرتخیلی را به وجود می-آورد.</p>				
---	--	--	--	--

## فصل سوم: روش تحقیق

## مقدمه

بسیاری از شرکت‌ها معتقدند نداشتن یک استراتژی مناسب در حوزه کلان داده منجر به عقب ماندن آن‌ها از بازار یا حتی شکست آن‌ها خواهد شد. اطلاعات، جزئی حیاتی از مدیریت کسب‌وکار است. باوجود کلان داده مدیران می‌توانند رفتار مشتریان را ارزیابی کنند، اطلاعات زیادی از کسب‌وکار خود به دست آورند و با استفاده از این دانش تصمیم‌گیری و عملکردشان را بهبود دهند. در عصر کلان داده‌ها می‌توانیم درک عمیق‌تری از بازار و مشتریان داشته باشیم تا جایکه ذائقه مشتریان در پایین‌ترین سطح را شناخته و بر اساس آن محصولات و خدمات جدیدی را تولید یا سفارشی سازی کنیم. بنابراین تجزیه و تحلیل کاربرد Big Data در فرآیند بازاریابی یکی از مهمترین مسائل در صنعت بازاریابی می‌باشد که ما در این فصل روش تحقیق موضوع را بررسی می‌کنیم.

### ۱-۳ نوع و ماهیت تحقیق

از نظر نوع، تحقیق حاضر یک تحقیق کاربردی است. پژوهش‌های کاربردی، پژوهش‌هایی هستند که با استفاده از نتایج تحقیقات بنیادی به منظور بهبود و به کمال رساندن رفتارها، روش‌ها، ابزارها، وسایل، تولیدات، ساختارها و الگوهای مورد استفاده جوامع انسانی انجام می‌شود. همچنین پژوهش‌های کاربردی پژوهش‌هایی هستند که نظریه‌ها، قانونمندی‌ها، اصول و فیزیکی که در تحقیقات پایه تدوین می‌شوند را برای حل مسائل اجرایی به کار می‌گیرد. با در نظر گرفتن این امر که نتایج حاصل از این تحقیق در تمامی صنعت‌ها و فعالیت‌های اقتصادی در هر زمینه‌ای قابل بهره‌برداری و استفاده می‌باشد این تحقیق کاربردی قلمداد می‌گردد. از جمله کاربردهای Big Data در صنعت بازاریابی می‌باشد با کمک Big Data می‌توان فهمید چه کسی، چه محصولی را در چه زمانی با چه قیمتی خریده است. و در نهایت به کمک Big Data می‌توان بین محصولی که مشتریان می‌بینند یا در موردش شنیده یا می‌خوانند، و چیزی که نهایتاً می‌خرند یا مصرف می‌کنند، ارتباطی منطقی پیدا کرد (سرمد، ۱۳۹۴).

نوع تحقیق و مطالعه کتابخانه‌ای و توصیفی-پیمایشی است. تحقیق توصیفی که گاهی به عنوان تحقیق غیرتجربی نامیده می‌شود، با روابط بین متغیرها، آزمودن فرضیه‌ها، پروراندن مفاهیم و قوانین کلی، اصول و یا نظریه‌هایی که دارای روایی جهان شمول است، سروکار دارد.

با توجه به مفاهیمی که در باب متغیرهای شناخته شده مرتبط با فرآیندهای بازاریابی ذکر شد، پژوهشگر با ذکر فرضیات خود قصد دارد خلا ذهنی خود را در باب رابطه‌ی متغیرهای مذکور با فرآیندها و تکنولوژی‌های مختلف برطرف کند. یکی از این فرآیندها که در پژوهش حاضر مدنظر واقع شده، کاربرد Big Data می‌باشد. داده‌های حجیم (Big Data)، به عنوان نوعی از داده‌ها که از آن برای تحلیل حجم‌های بیشتری از داده‌ها به کار برده می‌شود در فرضیات پژوهش به عنوان یکی از متغیرها گنجانده شده است. به عبارت دیگر، محقق با کشف رابطه‌ی Big Data و هریک از متغیرهای مرتبط با بازاریابی در پی این است که تا چه میزان Big Data، که ویژگی اصلی آن حجم گسترده و عظیمی از داده‌هاست، می‌تواند برای متغیرهای موردنظر کاربردی باشد. به عبارتی با پاسخ به فرضیات موردنظر می‌توان دریافت که آیا از Big Data برای مولفه‌های بازاریابی استفاده کرد یا خیر؟

فرضیه ۱: بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

فرضیه ۲: بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

فرضیه ۳: بین کاربرد Big Data و تامین نیازها و خواسته‌های مشتریان در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

فرضیه ۴: بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

فرضیه ۵: بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

## ۲-۳ روش گردآوری اطلاعات

در این پژوهش برای گردآوری اطلاعات از روش کتابخانه‌ای استفاده خواهد شد. اولین گام در راه دست یابی به اهداف پژوهش جمع‌آوری اطلاعات و در یافتن واقعیت‌ها است و در مورد روش کتابخانه‌ای این که روش کتابخانه‌ای در تمام تحقیقات علمی مورد استفاده قرار می‌گیرد. برای جمع‌آوری اطلاعات در زمینه‌های مبانی نظری و ادبیات تحقیق و پیشینه از کتاب‌ها، پایان نامه‌ها، مقالات و پایگاه‌های اطلاعاتی مرتبط با موضوع پژوهش (داخلی و خارجی) استخراج و استفاده خواهد شد. برای جمع‌آوری اطلاعات در زمینه‌های مبانی نظری و ادبیات تحقیق و پیشینه آن با استفاده از فیش برداری از کتب، مقالات، مجلات و استفاده از پایگاه‌های اطلاعاتی، اطلاعات مورد نظر جهت تکمیل سنجش چارچوب نظری پژوهش استفاده شده است (پرهیزگار، ۱۳۹۰). بعد از اتمام مطالعات کتابخانه‌ای مطالعه حاضر وارد مرحله میدانی می‌شود که



در حقیقت اقدام به جمع‌آوری اطلاعات به صورت میدانی شده و بر اساس متغیرهای تحقیق فرآیند تأثیر هر کدام از پارامترها سنجیده می‌شود.

### ۳-۳ روش کتابخانه‌ای (توصیفی - پیمایشی)

اولین گام در راه دست‌یابی به اهداف تحقیق جمع‌آوری اطلاعات و در یافتن واقعیت‌ها است و در مورد روش کتابخانه‌ای این که روش کتابخانه‌ای در تمام تحقیقات علمی مورد استفاده قرار می‌گیرد. برای جمع‌آوری اطلاعات در زمینه‌های مبانی نظری و ادبیات تحقیق و پیشینه از کتاب‌ها، پایان‌نامه‌ها، مقالات و پایگاه‌های اطلاعاتی مرتبط با موضوع پژوهش (داخلی و خارجی) استخراج و استفاده خواهد شد. برای جمع‌آوری اطلاعات در زمینه‌های مبانی نظری و ادبیات تحقیق و پیشینه آن با استفاده از فیش برداری از کتب، مقالات، مجلات و استفاده از پایگاه‌های اطلاعاتی، اطلاعات مورد نظر جهت تکمیل سنجش چارچوب نظری پژوهش استفاده شده است (پرهیزگار، ۱۳۹۰).

### ۳-۴ روش مطالعه میدانی

در مرحله دوم که تحقیق وارد مرحله میدانی می‌گردد، از پرسشنامه به عنوان ابزار گردآوری اطلاعات استفاده می‌گردد که با توجه به متغیرهای موضوع و رابطه‌ی مولفه‌های تحقیق اقدام به جمع‌آوری داده‌های مورد نیاز در تحقیق می‌شود. پرسشنامه مجموعه‌ای از سوال‌هاست که پاسخ دهنده با ملاحظه آنها پاسخ لازم را ارائه می‌دهد. این پاسخ‌ها داده مورد نیاز پژوهشگر را تشکیل می‌دهد.

### ۳-۵ ابزار گردآوری اطلاعات

ابزار گردآوری اطلاعات، وسایلی هستند که محقق به کمک آنها می‌تواند اطلاعات مورد نیاز را برای تجزیه و تحلیل و بررسی پدیده مورد مطالعه جمع‌آوری کند. ابزار که در این پژوهش استفاده خواهد شد، پرسشنامه است. پرسشنامه بر اساس فرضیات و متغیرهای تحقیق تهیه می‌شود. قبل از توزیع در بین نمونه مورد مطالعه ابتدا روایی و پایایی آن تعیین می‌شود.

### ۳-۶ پرسشنامه

پرسشنامه به عنوان یکی از متداولترین ابزار جمع‌آوری اطلاعات در تحقیقات پیمایشی، عبارت است از مجموعه‌ای از پرسش‌های هدف مدار که با بهره‌گیری از مقیاس‌های گوناگون، نظر، دیدگاه و بینش یک فرد پاسخگو را مورد سنجش قرار می‌دهد. پس از این که پرسشنامه مورد نظر تنظیم و افراد نمونه انتخابی نیز مشخص گردیدند نسبت به توزیع پرسشنامه اقدام می‌شود (دلاور، ۱۳۸۶).

### ۳-۹ روایی پرسشنامه

مقصود این است که آیا ابزار اندازه‌گیری مورد نظر می‌تواند ویژگی و خصوصیتی که ابزار برای آن طراحی شده است را اندازه‌گیری کند یا خیر؟ نظر کارشناسان و خبرگان می‌تواند کمک خوبی برای بهبود روایی ابزار اندازه‌گیری باشد. موضوع روایی از آن جهت اهمیت دارد که اندازه‌گیری‌های نامتناسب می‌تواند هر پژوهش علمی را بی ارزش سازد. برای محاسبه اعتبار پرسشنامه از روش اعتبار صوری استفاده شده است. به گونه ای که از نظرات متخصصان و صاحب‌نظران موضوع در رابطه با صحت و درستی پرسشنامه استفاده شده است (استراس و کورین، ۱۳۹۰).

### ۳-۱۰ پایایی پرسشنامه (قابلیت اعتماد)

پایایی که به دقت، اعتمادپذیری و ثبات نتایج آزمون اشاره می‌کند، از جمله ویژگی‌های فنی ابزار اندازه‌گیری بوده و با این امر سر و کار دارد که ابزار اندازه‌گیری در شرایط یکسان تا چه اندازه نتایج یکسانی بدست می‌دهد. همچنین، ضریب پایایی نشانه‌های مطلب است که تا چه اندازه ابزار اندازه‌گیری ویژگی‌های باثبات و یا ویژگی‌های متغیر آزمودنی را می‌سنجد.

قابلیت اعتماد که واژه‌هایی مانند پایایی، ثبات و اعتبار برای آن به کار برده می‌شود، یکی از ویژگی‌های ابزار اندازه‌گیری (پرسشنامه) است. ضریب قابلیت اعتماد نشانگر آن است که تا چه اندازه ابزار اندازه‌گیری ویژگی‌های با ثبات آزمودنی و یا ویژگی‌های متغیر و موقتی وی را می‌سنجد. برای اندازه‌گیری پایایی شاخصی به نام ضریب پایایی استفاده شد. دامنه ضریب پایایی از صفر تا ۱+ است. ضریب پایایی صفر معرف عدم پایایی و ضریب پایایی یک معرف پایایی کامل است. «پایایی کامل» واقعاً به ندرت دیده می‌شود و در صورت مشاهده قبل از هر چیز باید به نتایج حاصل شک کرد. برای محاسبه ضریب پایایی ابزار اندازه‌گیری، شیوه‌های مختلفی به کار برده می‌شود. از جمله: روش آلفای کرونباخ.

در روش آلفای کرونباخ بین نمره‌ی تک تک گویه‌ها با کل نمره‌ای ابزار یا آزمون همبستگی گرفته می‌شود و در آن فرض بر این است که انتظار می‌رود افراد که در گویه‌ی معینی نمره معینی را می‌گیرند در گویه‌های دیگر نیز آنگونه عمل نمایند، چون همه گویه‌های ابزار طراحی شده در راستای سنجش یک چیز یا خصیصه‌ی معین هستند (میرزایی، ۱۳۸۸).

در این تحقیق برای برآورد اعتبار پرسشنامه از روش آلفای کرونباخ استفاده شده است. برای بدست آوردن ضریب اعتبار پرسشنامه، ابتدا از طریق پیش آزمون ۲۰ پرسشنامه بین یک نمونه تصادفی توزیع شده سپس با استفاده از رایانه و نرم افزار SPSS ضریب آلفای کرونباخ محاسبه شده است.

(۳-۱)

$$r_a = \frac{j}{j-1} \left( 1 - \frac{\sum s_j^2}{s^2} \right)$$

j = تعداد زیر مجموعه سوال‌های پرسشنامه یا آزمون

$$s_j^2 = \text{واریانس زیر آزمون } j$$

$$s^2 = \text{واریانس کل آزمون.}$$

مقدار صفر این ضریب نشان دهنده عدم قابلیت اعتماد و ۱+ نشان دهنده قابلیت اعتماد کامل است.

زمانی یک پرسش نامه پایاست که مقدار آلفای کرونباخ بزرگتر از ۰/۷ باشد، و هر چه این مقدار به عدد ۱ نزدیک تر باشد پرسشنامه از پایداری بالاتری برخوردار است. میزان آلفای کرونباخ محاسبه شده برای پژوهش حاضر مساوی ۰/۸۳ که چون به عدد یک نزدیکتر است می توان گفت پرسشنامه از اعتبار بالایی برخوردار است.

## ۱۲-۳ جامعه آماری

جامعه‌ی آماری این پژوهش را مدیران و کارشناسان بازاریابی و کارمندان شرکت کسب و کار توسعه منابع انرژی توان می- باشند که تعداد آنها ۶۰ نفر به عنوان جامعه آماری می‌باشند.

## ۱۳-۳ روش نمونه‌گیری

باید توجه داشت که تعریف صحیح از جامعه دارای اهمیت زیادی است. به عبارتی تعریف جامعه باید چنان بیان شود که از نقطه نظر زمانی و مکانی همه واحدهای مورد مطالعه را در برگیرد و در ضمن با توجه به آن از شمول واحدهایی که نباید به مطالعه آنها پرداخته شود جلوگیری به عمل آید (حافظ نیا، ۱۳۸۹).

روش نمونه‌گیری در این تحقیق از نوع تصادفی ساده می‌باشد.

### ۱۴- روش تعیین حجم نمونه

حجم جامعه‌ی آماری این پژوهش ۶۰ نفر می‌باشند که برای به دست آوردن حجم نمونه از فرمول کوکران استفاده می‌شود.

(۳-۲)

$$n = \frac{\frac{z^2 pq}{d^2}}{1 + \frac{1}{N} \left( \frac{z^2 pq}{d^2} - 1 \right)}$$

N = حجم جامعه آماری

n = حجم نمونه

Z = مقدار متغیر نرمال واحد استاندارد، که در سطح اطمینان ۹۵ درصد برابر ۱/۹۶ می‌باشد.

p = مقدار نسبت صفت موجود در جامعه است. اگر در اختیار نباشد می‌توان آن را ۰/۵ در نظر گرفت.

q=(1-p) نسبتی از جمعیت فاقد صفت معین

d = درجه اطمینان یا دقت احتمالی مطلوب

$$\frac{\frac{1.96^2 \times 0.5 \times 0.5}{0.5^2}}{1 + \frac{1}{60} \left( \frac{1.96^2 \times 0.5 \times 0.5}{0.5} - 1 \right)} = ۶۰$$

حجم نمونه پژوهش فوق با استفاده از فرمول کوکران ۶۰ نفر تعیین شد.

### ۱۵-۳ تجزیه و تحلیل اطلاعات

با توجه به اینکه فرضیات ما رابطه ی بین دو متغیر را می‌سنجند، باید از آزمون آماری استفاده شود که میزان همبستگی و یا رابطه‌ی بین دو متغیر را نشان دهد که ضریب همبستگی پیرسون به عنوان یکی از آزمون‌های پارامتریک برای تجزیه و تحلیل این پژوهش استفاده شده‌است.

ضریب همبستگی همیشه عددی بین ۱ تا ۱- است. ضریب همبستگی بین ۰ تا ۱ به معنی داشتن همبستگی مثبت است و هرچه این ضریب به ۱ نزدیک‌تر باشد همبستگی قوی‌تر است. همبستگی مثبت یعنی با افزایش نمره یک متغیر نمره متغیر دیگر نیز افزایش می‌یابد..

### ۱۶-۳ نحوه تفسیر ضریب همبستگی

- ضریب بین ۰ تا ۰,۲۹ نشان دهنده همبستگی ضعیف

- ضریب بین ۰,۳۰ تا ۰,۶۹ نشان دهنده همبستگی متوسط

- ضریب بین ۰,۷۰ تا ۱ نشان دهنده همبستگی قوی

درمباحث آماری، ضریب همبستگی پیرسون یا ضریب همبستگی حاصل ضرب-گشتاور

پیرسون میزان همبستگی خطی بین دو متغیر تصادفی را می‌سنجد. مقدار این ضریب بین ۱- تا ۱ تغییر می‌کند که «۱» به معنای همبستگی مثبت کامل، «۰» به معنی نبود همبستگی، و «۱-» به معنی همبستگی منفی کامل است. این ضریب که کاربرد فراوانی در آمار دارد، توسط کارل پیرسون براساس ایده اولیه فرانسیس گالتون تدوین شد.

مدل مفهومی تحقیق



### جمع‌بندی فصل سوم

طی این پژوهش، با بررسی مکتوبات و اسناد به دست آمده از پژوهش‌های سایر محققان در زمینه‌های مشابه و بعضاً متفاوت، با موضوع تحقیق مورد بررسی شد. بر این اساس، طی این فصل در قالب روش‌شناسی و زیر شاخه آن یعنی روش کتابخانه‌ای، به تحقیق و بررسی روش‌های عنوان شده، پرداخته شد. سپس در ادامه برای انجام این پژوهش و جمع‌آوری اطلاعات و داده‌های مورد نیاز، از طریق سایت‌های معتبر اقدام شد. در انتها نیز در خصوص روش تجزیه و تحلیل پژوهش عنوان شد.

## فصل چهارم: تجزیه و تحلیل داده ها

## ۴-۱ مقدمه

داده های جمع آوری شده اعداد و ارقامی بدون معنی می باشند که از آمار برای معنی دار کردن آنها به منظور تحقق اهداف پژوهش ها و تحقیقات کمک گرفته می شود. تجزیه و تحلیل اطلاعات به عنوان بخشی از فرایند روش تحقیق علمی یکی از پایه های اصلی هر مطالعه و پژوهش به شمار می رود که بوسیله آن کلیه فعالیت های تحقیقی تا رسیدن به یک نتیجه، کنترل و هدایت می شوند.

تجزیه و تحلیل اطلاعات به عنوان بخشی از فرایند روش تحقیق علمی یکی از پایه های اصلی هر مطالعه و پژوهش به شمار می رود که بوسیله آن کلیه فعالیت های تحقیقی تا رسیدن به یک نتیجه، کنترل و هدایت می شوند. به عبارتی در این بخش، پژوهشگر برای پاسخگویی به مسأله تدوین شده یا تصمیم گیری در مورد رد یا تایید فرضیه یا فرضیاتی که برای تحقیق در نظر گرفته است از روش های مختلف تجزیه و تحلیل استفاده می کند. لذا ذکر این نکته ضروری است که تجزیه و تحلیل داده های بدست آمده به تنهایی برای یافتن پاسخ پرسش های پژوهش کافی نیست، تعبیر و تفسیر این داده ها نیز لازم است. ابتدا باید داده ها را تجزیه و تحلیل نمود و سپس نتایج این تجزیه و تحلیل را مورد تعبیر و تفسیر قرار داد. در این بخش به منظور توصیف ویژگی های نمونه، ابتدا داده های جمع آوری شده با استفاده از شاخص های آمار توصیفی خلاصه و طبقه بندی می شود، و سپس با استفاده از شاخص های آمار استنباطی به تایید یا رد فرضیات پرداخته می شود. به منظور اجرای شیوه پیمایشی، بر اساس مطالعات صورت گرفته، گویه های ابزار پرسشنامه تدوین شد و پس از بررسی اساتید، پرسشنامه نهایی و تصحیح شد. پس از سنجش اعتبار پرسشنامه، از میان جامعه آماری تحقیق، با استفاده از روش های تعیین حجم نمونه، حجم نمونه آماری تحقیق حاضر محاسبه گردید. سپس با استفاده از روش نمونه گیری به شیوه تصادفی اقدام به توزیع پرسشنامه در میان جامعه آماری شده است. پس از گردآوری پرسشنامه های تکمیل شده، اقدام به نمره دهی پاسخ ایشان شد. پس از نمره دهی هر گویه، به منظور تجزیه و تحلیل داده های آماری از نرم افزار SPSS و آزمون های آماری مناسب استفاده شد. سپس در مرحله تدوین فصل چهارم اقدام به توضیح آزمون ها و نتایج حاصله در قالب جداول و نمودارهایی شد. در نهایت بر اساس نتایج مطالعه حاضر و نیز مطالعات دیگران، پیشنهاداتی ارائه شد.

## ۴-۲ تحلیل های آماری توصیفی

آمار توصیفی منظم و طبقه بندی داده ها، نمایش ترسیمی، و محاسبه مقادیری از قبیل نما، میانگین، میانه و ... می باشد که حاکی از مشخصات یکایک اعضای جامعه مورد بحث است. در آمار توصیفی اطلاعات حاصل از یک گروه، همان گروه را توصیف می کند و اطلاعات به دست آمده به دسته جات مشابه تعمیم داده نمی شود. به طور کلی از سه روش در آمار توصیفی برای خلاصه سازی داده ها استفاده می شود:



➤ استفاده از جداول

➤ استفاده از نمودار

➤ محاسبه مقادیری خاص که نشان‌دهنده خصوصیات مهمی از داده‌ها باشند.

در این بخش یافته‌های توصیفی و مشخصات داده‌های جمع‌آوری شده از این پژوهش با استفاده از شاخص‌های آماری توصیفی بیان می‌شود.

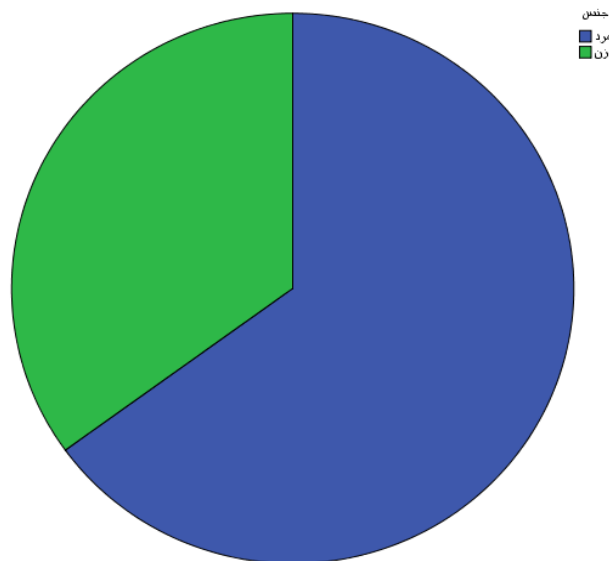
#### ۴-۲-۱ توزیع فراوانی بر حسب جنس

جامعه آماری موجود در این مطالعه مدیران و کارشناسان بازاریابی و کارمندان یک شرکت کسب و کار می باشد که جهت پرسش استفاده گردیده است. فراوانی و درصد جنسی پاسخگویان در جدول ۴-۱ ارائه شده است:

جدول ۴-۱ توزیع فراوانی بر حسب جنس

جنس	فراوانی	درصد فراوانی	درصد تجمعی
مرد	۳۹	۶۵	۶۵
زن	۲۱	۳۵	۱۰۰
کل	۶۰	۱۰۰	

همانطور که در جدول ۴-۱ نشان داده شده است، ۶۵ درصد (۳۹ نفر) از پاسخگویان را مرد و ۳۵ درصد (۲۱ نفر) از پاسخگویان را زن تشکیل می دهند.



نمودار ۴-۱ نمودار توزیع فراوانی بر حسب جنس

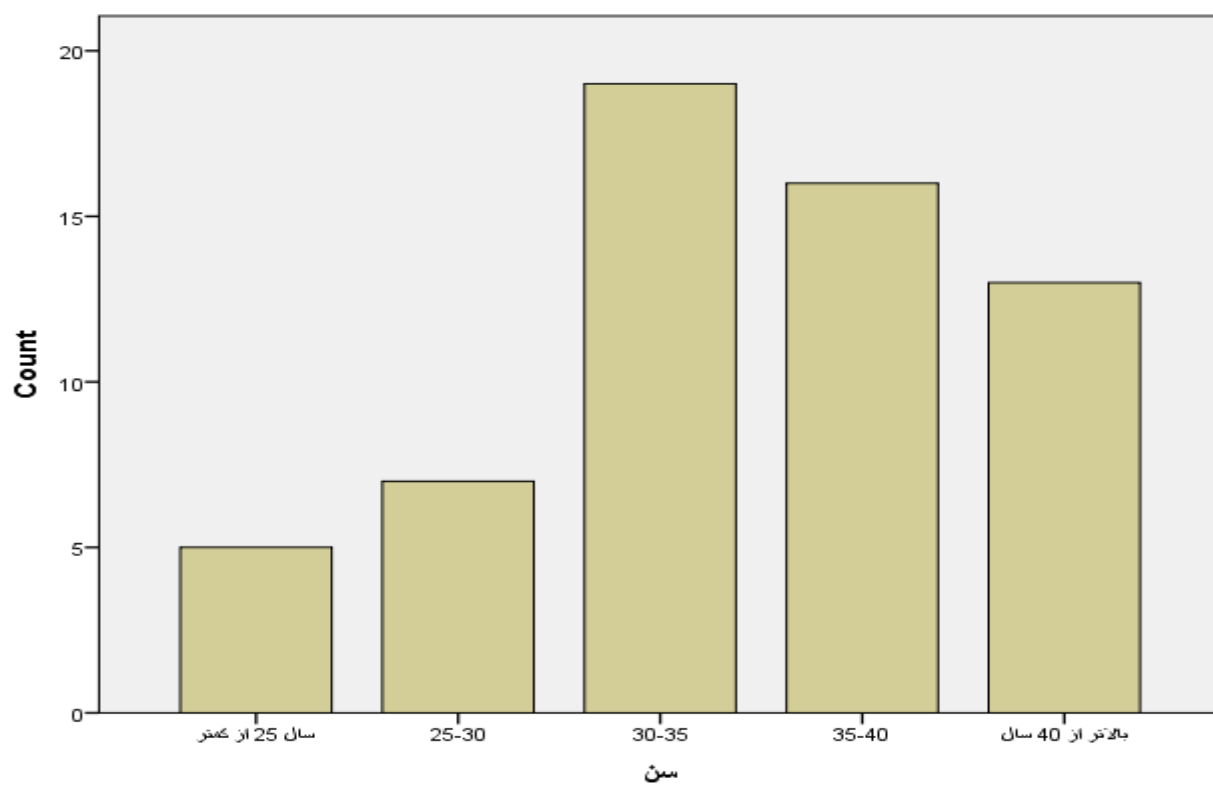
#### ۴-۲-۲ توزیع فراوانی بر حسب سن

بر اساس یافته های جدول ۴-۲ ملاحظه می گردد ۸,۳ درصد (۵ نفر) در گروه سنی کمتر از ۲۵ سال، ۱۱,۷ درصد (۷ نفر) در گروه سنی ۲۵-۳۰ سال، ۳۱,۷ درصد (۱۹ نفر) در گروه سنی ۳۰-۳۵ سال، ۲۶,۷ درصد (۱۶ نفر) در گروه سنی ۳۵-۴۰ سال و ۲۱,۷ درصد (۱۳ نفر) در گروه سنی بالاتر از ۴۰ سال می باشند که بیشترین تعداد افراد با ۳۱,۷ درصد متعلق به گروه سنی ۳۰-۳۵ سال می باشند. فراوانی و درصد رده های سنی افراد در جدول ۴-۲ ارائه شده است:

جدول ۴-۲ توزیع فراوانی بر حسب سن

سن	فراوانی	درصد فراوانی	درصد تجمعی
کمتر از ۲۵ سال	۵	۸,۳	۸,۳
۲۵-۳۰ سال	۷	۱۱,۷	۲۰
۳۰-۳۵ سال	۱۹	۳۱,۷	۵۱,۷

۷۸,۳	۲۶,۷	۱۶	۳۵-۴۰ سال
۱۰۰	۲۱,۷	۱۳	بالاتر از ۴۰ سال
	۱۰۰	۶۰	کل



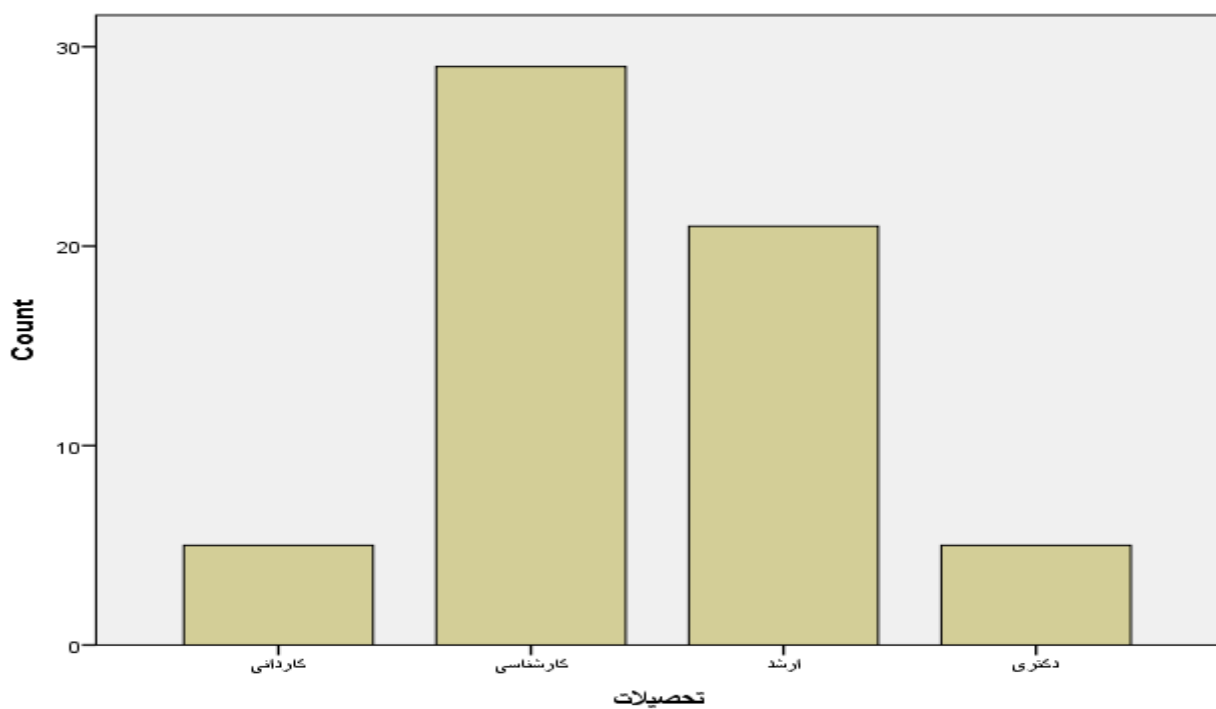
نمودار ۲-۴ نمودار توزیع فراوانی بر حسب سن

#### ۳-۲-۴ توزیع فراوانی بر حسب تحصیلات

بر اساس یافته های جدول ۳-۴ ملاحظه می گردد ۸,۳ درصد (۵ نفر) کاردانی، ۴۸,۴ درصد (۲۹ نفر) لیسانس و ۳۵ درصد (۲۱ نفر) فوق لیسانس و ۸,۳ درصد (۵ نفر) دکتری می باشند که بیشترین تعداد افراد با ۴۴,۸ درصد در مقطع کارشناسی می باشند. فراوانی و درصد سطح تحصیلات افراد در جدول ۳-۴ ارائه شده است:

جدول ۳-۴ توزیع فراوانی بر حسب تحصیلات

تحصیلات	فراوانی	درصد فراوانی	درصد تجمعی
کاردانی	۵	۸,۳	۸,۳
کارشناسی	۲۹	۴۸,۴	۵۶,۷
کارشناسی ارشد	۲۱	۳۵	۹۱,۷
دکتری	۵	۸,۳	۱۰۰
کل	۶۰	۱۰۰	



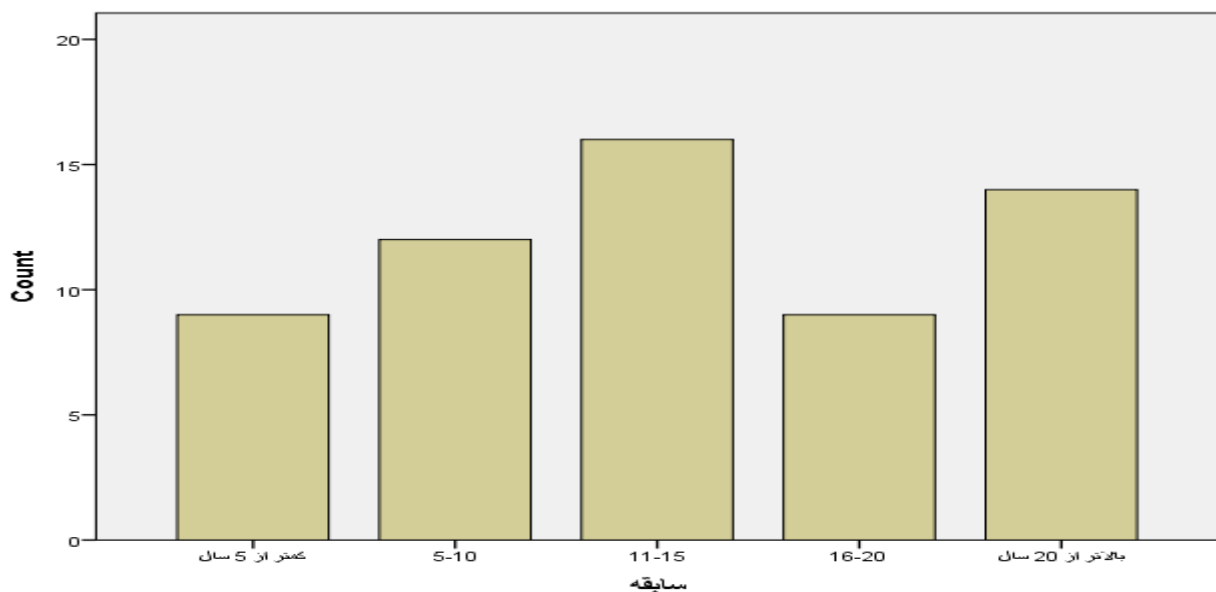
نمودار ۳-۴ نمودار توزیع فراوانی بر حسب تحصیلات

#### ۵-۲-۴ توزیع فراوانی بر حسب سابقه خدمت

بر اساس یافته های جدول ۵-۴ ملاحظه می گردد ۱۵ درصد (۹ نفر) کمتر از ۵ سال، ۲۰ درصد (۱۲ نفر) ۵-۱۰ سال، ۲۶,۷ درصد (۱۶ نفر) ۱۰-۱۵ سال، ۱۵ درصد (۹ نفر) ۱۵-۲۰ سال و ۲۳,۳ درصد (۱۴ نفر) بیش از ۲۰ سال دارای سابقه خدمت می باشند. فراوانی و درصد رده های مختلف سابقه خدمت افراد در جدول ۵-۴ ارائه شده است:

جدول ۵-۴ توزیع فراوانی بر حسب سابقه خدمت

سابقه خدمت	فراوانی	درصد فراوانی	درصد تجمعی
کمتر از ۵ سال	۹	۱۵	۱۵
۵-۱۰ سال	۱۲	۲۰	۳۵
۱۰-۱۵ سال	۱۶	۲۶,۷	۶۱,۷
۱۵-۲۰ سال	۹	۱۵	۷۶,۷
بیش از ۲۰ سال	۱۴	۲۳,۳	۱۰۰
کل	۶۰	۱۰۰	



نمودار ۴-۵ نمودار توزیع فراوانی بر حسب سابقه خدمت

#### ۴-۳- توصیف متغیرهای تحقیق

در این بخش شاخص های آماری میانه، انحراف استاندارد و درصد خطا ( متغیر مستقل و وابسته) گزارش شده است. در واقع این شاخص ها می تواند به درک بیش تر متغیرهای پژوهش کمک کند.

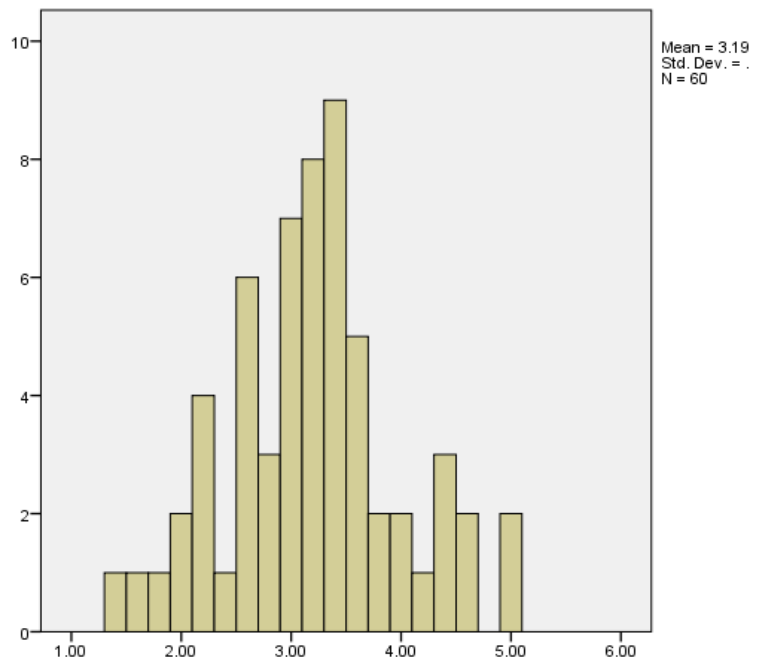
میانگین از متداولترین معیارهای گرایش به مرکز هستند. مطالعه توزیع یک جامعه آماری مقدار نماینده که اندازه ها در اطراف آن توزیع شده اند را مقدار مرکزی می نامند و هر معیار عددی را که معرف مرکز مجموعه داده ها باشد، معیار گرایش به مرکز می نامند. یکی از شاخص های پراکندگی، انحراف معیار است که نشان می دهد به طور میانگین داده ها چه مقدار از مقدار متوسط فاصله دارند. اگر انحراف معیار مجموعه ای از داده ها نزدیک به صفر باشد، نشانه آن است که داده ها نزدیک به میانگین هستند و پراکندگی اندکی دارند؛ در حالی که انحراف معیار بزرگ بیانگر پراکندگی قابل توجه داده ها می باشد. در واقع انحراف معیار برای تعیین ضریب اطمینان در تحلیل های آماری نیز به کار می رود. در مطالعات علمی، معمولاً داده های با انحراف معیار بیشتر از دو به عنوان داده های پرت در نظر گرفته و از تحلیل، خارج می شوند.

#### ۴-۳-۱ یکپارچگی

باتوجه به جدول (۴-۶) مشاهده می شود که کمترین مقدار متغیر "یکپارچگی" ۱ و بالاترین مقدار آن ۵، میانگین ۳,۰۳۷۵ و انحراف معیار ۰,۹۰۱۷۷ و واریانس آن ۰,۸۱۳ است. همچنین از آنجایی که مقدار چولگی در بازه ی (-۳ و +۳) قرار دارد، بنابراین می توان گفت داده ها از توزیع نرمال برخوردار هستند.

جدول ۴-۶) توصیف متغیر یکپارچگی

یکپارچگی	تعداد	حداقل	حداکثر	میانگین	انحراف استاندارد	معیار	واریانس	کشیدگی	چولگی
یکپارچگی	۶۰	۱	۵	۳,۱۹۰۰	۰,۷۹۳۷۰		۰,۶۲۸	۰,۱۴۵	۰,۰۵۳



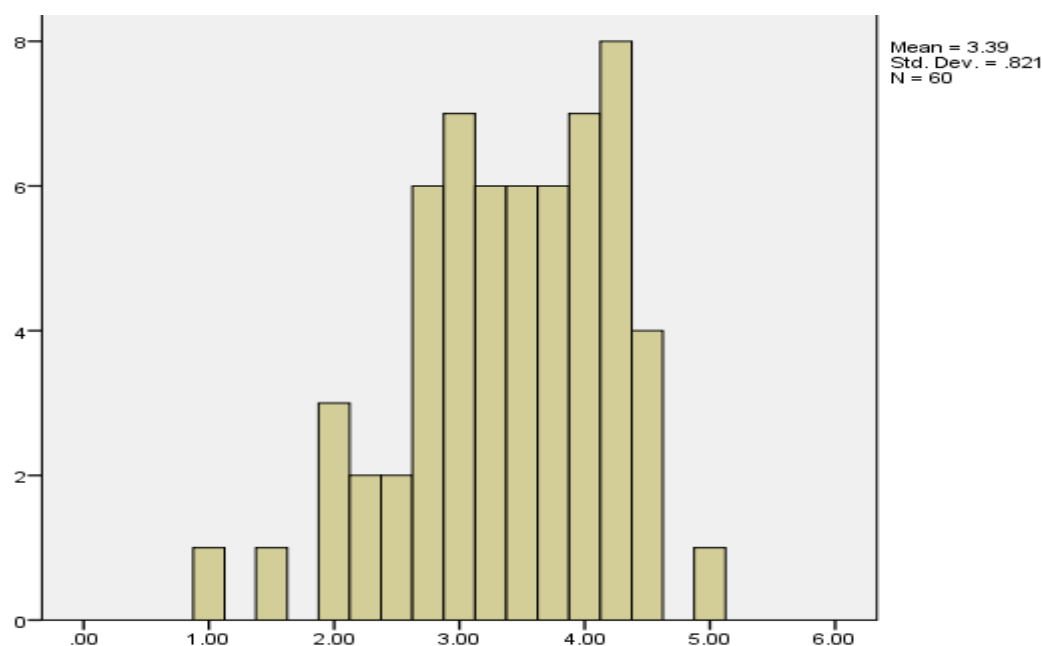
نمودار ۴-۶ هیستوگرام متغیر یکپارچگی

#### ۴-۳-۲ جذب مشتریان

باتوجه به جدول (۴-۷) مشاهده می شود که کمترین مقدار متغیر "جذب مشتریان" ۲ و بالاترین مقدار آن ۵، میانگین ۳,۴۷۷۸ و انحراف معیار ۰,۷۴۷۵۴ و واریانس آن ۰,۵۵۹ است. همچنین از آنجایی که مقدار چولگی در بازه ی (۳- و ۳+) قرار دارد، بنابراین می توان گفت داده ها از توزیع نرمال برخوردار هستند.

جدول ۴-۷) توصیف متغیر جذب مشتریان

	تعداد	حداقل	حداکثر	میانگین	انحراف استاندارد	معیار واریانس	کشیدگی	چولگی
جذب مشتریان	۶۰	۱	۵	۳,۳۹۱۷	۰,۸۲۱۲۸	۰,۶۷۵	-۰,۵۸۳	۰,۱۷۲



نمودار ۴-۷ هیستوگرام متغیر جذب مشتریان

#### ۴-۳-۳ تامین نیازها و خواسته های مشتریان

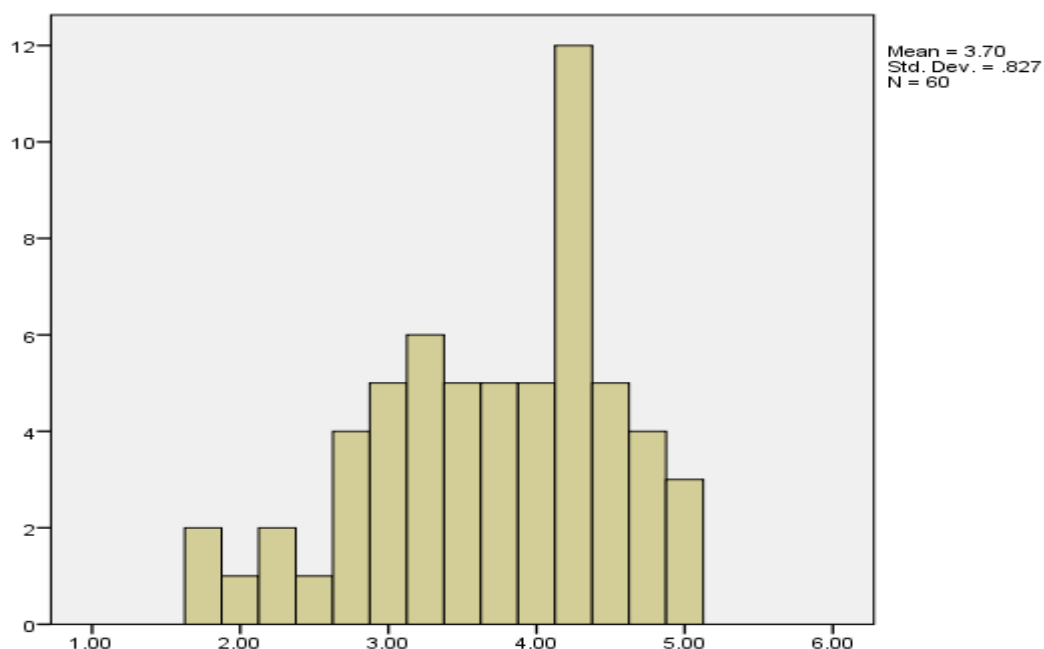
باتوجه به جدول (۴-۸) مشاهده می شود که کمترین مقدار متغیر "تامین نیازها و خواسته های مشتریان" ۱ و بالاترین مقدار آن ۵، میانگین ۳,۴۰۰ و انحراف معیار ۰,۸۷۸۲۸ و واریانس آن ۰,۷۷۱ است. همچنین از آنجایی که مقدار چولگی در بازه ی (-۳ و +۳) قرار دارد، بنابراین می توان گفت داده ها از توزیع نرمال برخوردار هستند.

#### جدول ۴-۸ توصیف متغیر تامین نیازها و خواسته های مشتریان

	تعداد	حداقل	حداکثر	میانگین	انحراف	معیار	واریانس	کشیدگی	چولگی
--	-------	-------	--------	---------	--------	-------	---------	--------	-------



			استاندارد					
تامین نیازها و خواسته های مشتریان	۶۰	۱	۵	۳,۶۹۵۸	۰,۸۲۷۳۵	۰,۶۸۵	-۰,۵۲۳	-۰,۳۸۶



نمودار ۴-۸ هیستوگرام متغیر تامین نیازها و خواسته های مشتریان

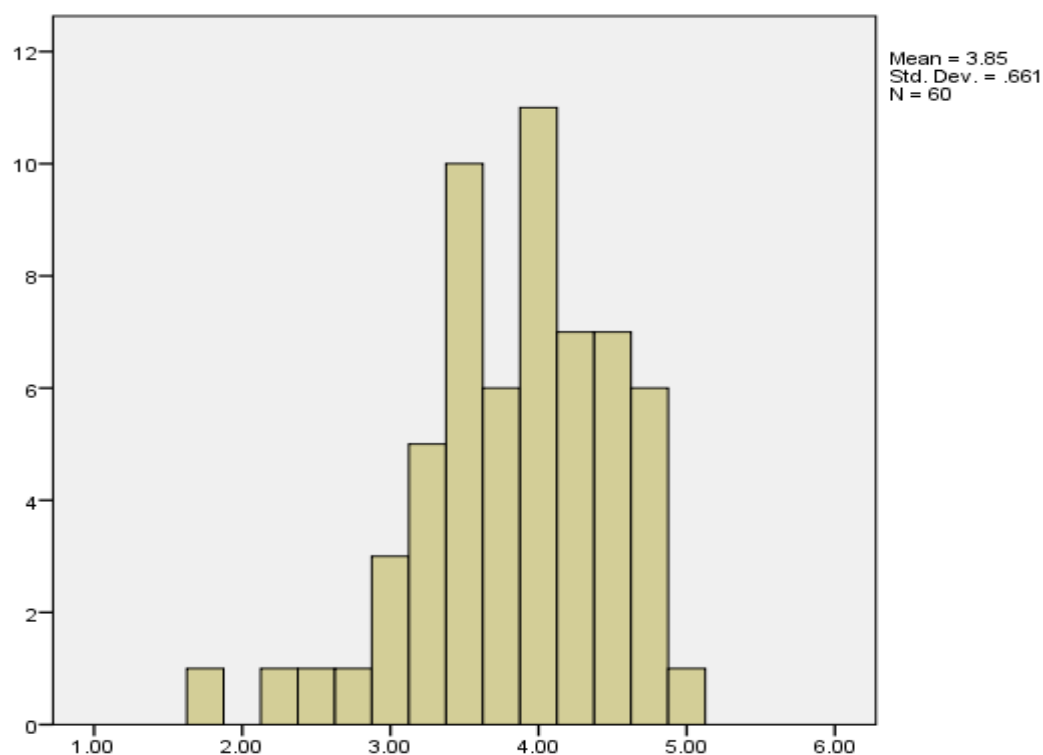
#### ۴-۳-۴ دسترس پذیری

باتوجه به جدول (۹-۴) مشاهده می شود که کمترین مقدار متغیر "دسترس پذیری" ۱ و بالاترین مقدار آن ۵، میانگین ۳,۶۰۰ و انحراف معیار ۰,۸۲۲۹۳ و واریانس آن ۰,۶۷۷ است. همچنین از آنجایی که مقدار چولگی در بازه ی (۳- و ۳+) قرار دارد، بنابراین می توان گفت داده ها از توزیع نرمال برخوردار هستند.

#### جدول ۹-۴ توصیف متغیر تامین دسترس پذیری

تعداد	حداقل	حداکثر	میانگین	انحراف استاندارد	معیار	واریانس	کشیدگی	چولگی
-------	-------	--------	---------	------------------	-------	---------	--------	-------

دسترس پذیری	۶۰	۱	۵	۳,۸۴۵۸	۰,۶۶۰۷۸	۰,۴۳۷	-۰,۷۲۲	۰,۷۸۵
-------------	----	---	---	--------	---------	-------	--------	-------



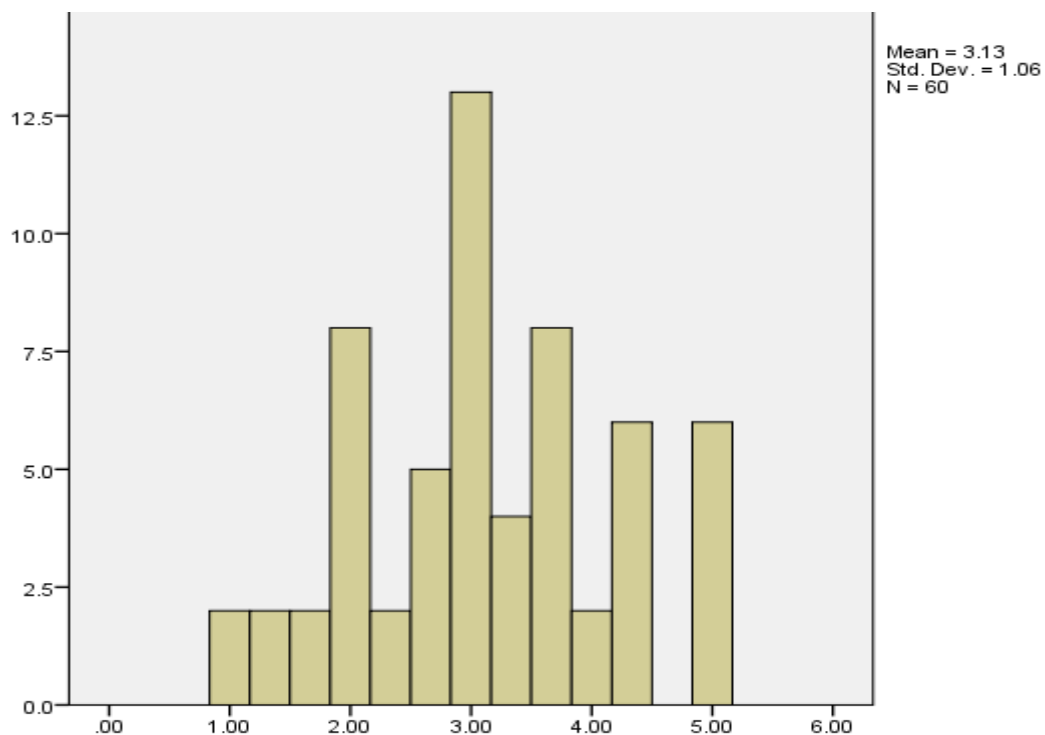
نمودار ۹-۴ هیستوگرام متغیر دسترس پذیری

#### ۵-۳-۴ مدیریت پایگاه داده

باتوجه به جدول (۱۰-۴) مشاهده می شود که کمترین مقدار متغیر "مدیریت پایگاه داده" ۱ و بالاترین مقدار آن ۵، میانگین ۴,۰۰۵۶ و انحراف معیار ۰,۸۸۶۱۶ و واریانس آن ۰,۷۸۵ است. همچنین از آنجایی که مقدار چولگی در بازه ی (۳- و ۳+) قرار دارد، بنابراین می توان گفت داده ها از توزیع نرمال برخوردار هستند.

#### جدول ۱۰-۴ توصیف متغیر مدیریت پایگاه داده

مدیریت پایگاه داده	تعداد	حداقل	حداکثر	میانگین	انحراف استاندارد	معیار	واریانس	کشیدگی	چولگی
	۶۰	۱	۵	۳,۱۲۷۸	۱,۰۵۹۶۰		۱,۱۲۳	۰,۰۳۶	-۰,۵۸۳



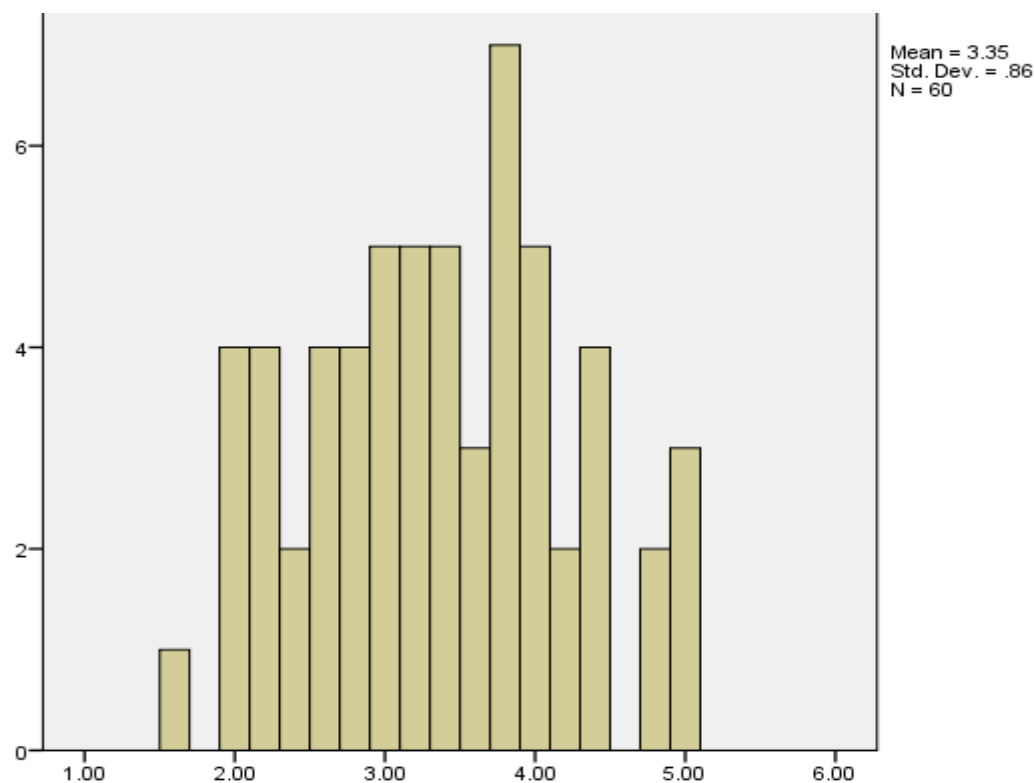
نمودار ۴-۱ هیستوگرام متغیر مدیریت پایگاه داده

#### ۴-۳-۶ کاربرد Big Data

باتوجه به جدول (۴-۱۱) مشاهده می شود که کمترین مقدار متغیر " کاربرد Big Data " ۱ و بالاترین مقدار آن ۵ ، میانگین ۴,۰۰۵۶ و انحراف معیار ۰,۸۸۶۱۶ و واریانس آن ۰,۷۸۵ است. همچنین از آنجایی که مقدار چولگی در بازه ی (۳- و ۳+) قرار دارد، بنابراین می توان گفت داده ها از توزیع نرمال برخوردار هستند.

جدول ۴-۱۱) توصیف متغیر کاربرد Big Data

چولگی	کشیدگی	واریانس	معیار انحراف استاندارد	میانگین	حداکثر	حداقل	تعداد	کاربرد Big Data
-۰,۷۱۱	۰,۰۶۵	۰,۷۳۹	۰,۸۵۹۵۴	۳,۳۴۶۷	۵	۱	۶۰	



نمودار ۴-۱۱ هیستوگرام متغیر کاربرد Big Data

#### ۴-۴ تجزیه و تحلیل استنباطی داده ها

در آزمون فرضیه‌ها یا آزمون معناداری، پژوهشگر، فرضیه صفر را رد یا قبول می‌کند. یعنی اینکه اگر  $H_0$  پذیرفته شود، فرض می‌شود که  $H_1$  رد شده است و اگر  $H_0$  رد شود  $H_1$  پذیرفته شود. برای تعیین معناداری آماری یک مطالعه پژوهشی، پژوهشگر باید سطح احتمالی یا سطح معناداری آن را تعیین کند؛ تا فرضیه صفر در مقابل آن آزموده شود. اگر

نتایج مطالعه احتمالی، کمتر از این سطح را نشان دهد، پژوهشگر می‌تواند فرضیه صفر را رد کند. هرگاه اثبات نتیجه ی پژوهش احتمال بالایی داشته باشد، پژوهشگر باید فرضیه صفر را تأیید کند؛ در حقیقت، از آنجا که فرضیه صفر، عموماً بیان نمی‌شود، قبول یا رد، در فرضیه "پژوهش" به کار بسته می‌شود، نه در فرضیه صفر.

به منظور تدوین فرضیه‌های پژوهش، پژوهشگر هم از قیاس و هم از استقرا استفاده می‌کند. بنابراین برای تنظیم فرضیه دو منبع بالقوه وجود دارد:

الف) تئوری‌های کلی و کامل و جامع موجود؛ این فرضیه‌ها با استفاده از یک رشته قیاس‌های معین منجر به پیش‌بینی‌هایی می‌شود که در صورت شرایط معین، نتایج مشخصی به‌دست خواهد آمد.

ب) نتایج پژوهش‌های انجام شده پیشین؛ هدف این فرضیه‌ها استقرا رابطه بین دو یا چند متغیر در زمان حال است.

#### ۴-۵ بررسی طبیعی بودن توزیع داده‌ها

برای استفاده از تکنیک‌های آماری ابتدا باید مشخص شود که داده‌های جمع‌آوری شده از توزیع نرمال برخوردار است یا خیر؟ زیرا در صورت نرمال بودن توزیع داده‌های جمع‌آوری شده برای آزمون فرضیه‌ها می‌توان از آزمون‌های پارامتریک استفاده نمود و در صورت غیر نرمال بودن از آزمون‌های ناپارامتریک استفاده کرد. برای این منظور در این پژوهش از آزمون معتبر کلموگروف-اسمیرنوف جهت بررسی نرمال بودن توزیع متغیرهای اصلی استفاده می‌شود. این آزمون در حالت تک نمونه‌ای به مقایسه تابع توزیع تجمعی مشاهده شده با تابع توزیع تجمعی مورد انتظار در یک متغیر در سطح سنجش فاصله-ای می‌پردازد. در تفسیر نتایج آزمون، چنانچه مقدار سطح خطای مشاهده شده از  $0.05$  بیشتر باشد، در آن صورت توزیع مشاهده شده با توزیع نظری یکسان است و تفاوتی بین این دو وجود ندارد. یعنی توزیع بدست آمده توزیع نرمال است. اما چنانچه مقدار معنی داری از  $0.05$  کوچکتر باشد آنگاه توزیع مشاهده شده با توزیع مورد انتظار متفاوت است و توزیع فوق نرمال نخواهد بود. این آزمون با توجه به فرضیات زیر به بررسی نرمال بودن داده می‌پردازد.

$H_0$ : بین فراوانی‌های مشاهده شده و مورد انتظار تفاوت وجود ندارد (توزیع نرمال هست).

$H_1$ : بین فراوانی‌های مشاهده شده و مورد انتظار تفاوت وجود دارد (توزیع نرمال نیست).

جدول ۴-۱۲ تست نرمالیت متغیرها

متغیر	نوع توزیع به سطح	مقدار	تایید	نتیجه
-------	------------------	-------	-------	-------

کار گرفته معناداری خطا	فرضیه	کار شده	گرفته معناداری خطا	فرضیه
یکپارچگی	نرمال	۰,۱۱۲	۰/۰۵	$H_0$ نرمال
جذب مشتری	نرمال	۰,۱۰۴	۰/۰۵	$H_0$ نرمال
تامین نیازها و خواسته های مشتریان	نرمال	۰,۱۴۹	۰/۰۵	$H_0$ نرمال
دسترس پذیری	نرمال	۰,۱۲۶	۰/۰۵	$H_0$ نرمال
مدیریت پایگاه داده	نرمال	۰,۱۱۵	۰/۰۵	$H_0$ نرمال
کاربرد Big Data	نرمال	۰,۰۸۴	۰/۰۵	$H_0$ نرمال

با توجه به مقادیر حاصل از آماره اسمیرنوف - کلموگروف جدول ۴-۱۲، می توان استنباط نمود که توزیع مورد انتظار با توزیع مشاهده شده برای تمام متغیرها تفاوت معنی داری ندارد و بنابراین توزیع این متغیرها نرمال است.

### آزمون فرضیات

فرضیه ۱: بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

فرضیه ۲: بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

فرضیه ۳: بین کاربرد Big Data و تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

فرضیه ۴: بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

فرضیه ۵: بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

**فرضیه اول: بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد.**

$H_0$ : به نظر می رسد بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود ندارد.

$H_1$ : به نظر می رسد بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

جدول ۴-۱۳ ضریب همبستگی پیرسون فرضیه اول

		کاربرد Big Data		یکپارچگی
Big Data	کاربرد	سطح پیرسون	۱	۰,۷۲۲
		سطح معناداری		۰,۰۰۰
		تعداد	۶۰	۶۰
یکپارچگی		سطح پیرسون	۰,۷۲۲	۱
		سطح معناداری	۰,۰۰۰	
		تعداد	۶۰	۶۰

همانطور که در جدول ۴-۱۳ مشاهده می کنید، مقدار سطح معناداری فرضیه اول کمتر از ۰/۰۵ است؛ بنابراین بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد. بنابراین می توان گفت که فرضیه اول تایید و فرض صفر آن رد می شود. مقدار و علامت ضریب این آزمون نیز جهت و قدرت رابطه را نشان می دهد. از آنجایی که علامت ضریب مثبت است بنابراین رابطه مستقیم بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی وجود دارد. پس می توان گفت که بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

جدول ۴-۱۴ آزمون F (آنالیز واریانس) جهت معنادار بودن رگرسیون

متغیر	مجموع مربعات	میانگین مربعات	درجه آزادی	F	سطح معناداری
رگرسیون	۲۲,۶۹۷	۲۲,۶۹۷	۱	۶۳,۰۰۹	۰,۰۰۰
باقی مانده	۲۰,۸۹۲	۰,۳۶۰	۵۸		
کل	۴۳,۵۸۹		۵۹		

جدول ۴-۱۴ میزان متغیر مستقل یعنی یکپارچگی در فرآیندهای بازاریابی متغیر وابسته یعنی کاربرد Big Data را نشان می دهد. همانطور که مشاهده می کنید مقدار معنی داری کوچکتر از ۰/۰۵ است بنابراین به طور کلی می توان گفت متغیر مستقل بر متغیر وابسته اثرگذار است.

جدول ۴-۱۵ ضرایب تاثیر رگرسیون

متغیر	بتا	T	سطح معناداری
مقدار ثابت	-	۲,۶۲۷	۰,۰۱۱
یکپارچگی	۰,۷۲۲	۷,۹۳۸	۰,۰۰۰

مقدار ثابت همان عرض از مبدا است و میزان متغیر وابسته را بدون دخالت متغیر مستقل نشان می دهد. با توجه به نتایج جدول فوق می توان گفت با ارتقا یک واحد از متغیر مستقل به میزان ضریب نوشته شده متغیر وابسته ارتقا پیدا خواهد کرد. آماره  $t$  اهمیت نسبی متغیر مستقل را نشان می دهد. با توجه به مقدار این آماره و سطح خطای کمتر از ۰,۰۵ می توان گفت متغیر مورد نظر تاثیر آماری معنی داری در تبیین تغییرات متغیر وابسته داشته است. ضرایب رگرسیون نیز تاثیر متغیر مستقل بر متغیر وابسته را تایید می کند. با توجه به ضریب یکپارچگی، فرضیه پژوهش مورد تایید قرار می گیرد. در نتیجه می توان گفت بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد. یعنی فرض صفر فرضیه اول رد و فرض یک تایید می گردد.

**فرضیه دوم: بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.**

$H_0$ : به نظر می رسد بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی رابطه معناداری وجود ندارد.

$H_1$ : به نظر می رسد بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

جدول ۴-۱۶ ضریب همبستگی پیرسون فرضیه دوم



جذب مشتری		کاربرد Big Data	
کاربرد Big Data	سطح پیرسون	۱	۰,۳۵۴
	سطح معناداری		۰,۰۰۵
	تعداد	۶۰	۶۰
جذب مشتری	سطح پیرسون	۰,۳۵۴	۱
	سطح معناداری	۰,۰۰۵	
	تعداد	۶۰	۶۰

همانطور که در جدول ۴-۱۶ مشاهده می کنید، مقدار سطح معناداری فرضیه دوم کمتر از ۰/۰۵ است؛ بنابراین بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی رابطه معناداری وجود دارد. بنابراین می توان گفت که فرضیه دوم تایید و فرض صفر آن رد می شود. مقدار و علامت ضریب این آزمون نیز جهت و قدرت رابطه را نشان می دهد. از آنجایی که علامت ضریب مثبت است بنابراین رابطه مستقیم بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی وجود دارد. پس می توان گفت که بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

جدول ۴-۱۷ آزمون F (آنالیز واریانس) جهت معنادار بودن رگرسیون

متغیر	مجموع مربعات	میانگین مربعات	درجه آزادی	F	سطح معناداری
رگرسیون	۵,۴۶۹	۵,۴۶۹	۱	۸,۳۲۲	۰,۰۰۵
باقی مانده	۳۸,۱۲۰	۰,۶۵۷	۵۸		
کل	۴۳,۵۸۹		۵۹		

جدول ۴-۱۷ میزان متغیر مستقل یعنی جذب مشتری در فرآیندهای بازاریابی متغیر وابسته یعنی کاربرد Big Data را نشان می دهد. همانطور که مشاهده می کنید مقدار معنی داری کوچکتر از ۰/۰۵ است بنابراین به طور کلی می توان گفت متغیر مستقل بر متغیر وابسته اثرگذار است.

جدول ۴-۱۸ ضرایب تاثیر رگرسیون

متغیر	بتا	T	سطح معناداری
مقدار ثابت	-	۴,۶۶۱	۰,۰۰۰
جذب مشتری	۰,۳۵۴	۲,۸۸۵	۰,۰۰۵

مقدار ثابت همان عرض از مبدا است و میزان متغیر وابسته را بدون دخالت متغیر مستقل نشان می دهد. با توجه به نتایج جدول فوق می توان گفت با ارتقا یک واحد از متغیر مستقل به میزان ضریب نوشته شده متغیر وابسته ارتقا پیدا خواهد کرد. آماره  $t$  اهمیت نسبی متغیر مستقل را نشان می دهد. با توجه به مقدار این آماره و سطح خطای کمتر از ۰,۰۵ می توان گفت متغیر مورد نظر تاثیر آماری معنی داری در تبیین تغییرات متغیر وابسته داشته است. ضرایب رگرسیون نیز تاثیر متغیر مستقل بر متغیر وابسته را تایید می کند. با توجه به ضریب جذب مشتری، فرضیه پژوهش مورد تایید قرار می گیرد. در نتیجه می توان گفت بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی رابطه معناداری وجود دارد. یعنی فرض صفر فرضیه دوم رد و فرض یک تایید می گردد.

**فرضیه سوم: بین کاربرد Big Data و تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی رابطه معناداری وجود دارد.**

$H_0$ : به نظر می رسد بین کاربرد Big Data و تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی رابطه معناداری وجود ندارد.

$H_1$ : به نظر می رسد بین کاربرد Big Data و تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

جدول ۴-۱۹ ضریب همبستگی پیرسون فرضیه سوم

		کاربرد Big Data	تامین نیازها و خواسته های مشتریان
Big Data	سطح پیرسون	۱	۰,۴۱۳
	سطح معناداری		۰,۰۰۱
	تعداد	۶۰	۶۰
تامین نیازها و	سطح پیرسون	۰,۴۱۳	۱

های	خواسته	سطح معناداری	۰,۰۰۱	
	تعداد مشتریان		۶۰	۶۰

همانطور که در جدول ۴-۱۹ مشاهده می کنید، مقدار سطح معناداری فرضیه سوم کمتر از ۰/۰۵ است؛ بنابراین بین کاربرد Big Data و تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی رابطه معناداری وجود دارد. بنابراین می توان گفت که فرضیه سوم تایید و فرض صفر آن رد می شود. مقدار و علامت ضریب این آزمون نیز جهت و قدرت رابطه را نشان می دهد. از آنجایی که علامت ضریب مثبت است بنابراین رابطه مستقیم بین کاربرد Big Data و تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی وجود دارد. پس می توان گفت که بین کاربرد Big Data و تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

جدول ۴-۲۰ آزمون F (آنالیز واریانس) جهت معنادار بودن رگرسیون

متغیر	مجموع مربعات	میانگین مربعات	درجه آزادی	F	سطح معناداری
رگرسیون	۷,۴۳۴	۷,۴۳۴	۱	۱۱,۹۲۵	۰,۰۰۱
باقی مانده	۳۶,۱۵۶	۰,۶۲۳	۵۸		
کل	۴۳,۵۸۹		۵۹		

جدول ۴-۲۰ میزان متغیر مستقل یعنی تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی متغیر وابسته یعنی کاربرد Big Data را نشان می دهد. همانطور که مشاهده می کنید مقدار معنی داری کوچکتر از ۰/۰۵ است بنابراین به طور کلی می توان گفت متغیر مستقل بر متغیر وابسته اثرگذار است.

جدول ۴-۲۱ ضرایب تاثیر رگرسیون

متغیر	بتا	T	سطح معناداری
مقدار ثابت	-	۳,۷۴۴	۰,۰۰۰

مقدار ثابت همان عرض از مبدا است و میزان متغیر وابسته را بدون دخالت متغیر مستقل نشان می دهد. با توجه به نتایج جدول فوق می توان گفت با ارتقا یک واحد از متغیر مستقل به میزان ضریب نوشته شده متغیر وابسته ارتقا پیدا خواهد کرد. آماره  $t$  اهمیت نسبی متغیر مستقل را نشان می دهد. با توجه به مقدار این آماره و سطح خطای کمتر از ۰,۰۵ می توان گفت متغیر مورد نظر تاثیر آماری معنی داری در تبیین تغییرات متغیر وابسته داشته است. ضرایب رگرسیون نیز تاثیر متغیر مستقل بر متغیر وابسته را تایید می کند. با توجه به ضریب تامین نیازها و خواسته های مشتریان، فرضیه پژوهش مورد تایید قرار می گیرد. در نتیجه می توان گفت بین کاربرد Big Data و تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی رابطه معناداری وجود دارد. یعنی فرض صفر فرضیه سوم رد و فرض یک تایید می گردد.

#### فرضیه چهارم: بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

$H_0$ : به نظر می رسد بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی رابطه معناداری وجود ندارد.

$H_1$ : به نظر می رسد بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

جدول ۴-۲۲ ضریب همبستگی پیرسون فرضیه چهارم

		کاربرد Big Data	
		دسترس پذیری	
Big Data کاربرد	سطح پیرسون	۱	۰,۵۶۹
	سطح معناداری		۰,۰۰۰
	تعداد	۶۰	۶۰

دسترس پذیری	سطح پیرسون	۰,۵۶۹	۱
	سطح معناداری	۰,۰۰۰	
	تعداد	۶۰	۶۰

همانطور که در جدول ۴-۲۲ مشاهده می کنید، مقدار سطح معناداری فرضیه چهارم کمتر از ۰/۰۵ است؛ بنابراین بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی رابطه معناداری وجود دارد. بنابراین می توان گفت که فرضیه چهارم تایید و فرض صفر آن رد می شود. مقدار و علامت ضریب این آزمون نیز جهت و قدرت رابطه را نشان می دهد. از آنجایی که علامت ضریب مثبت است بنابراین رابطه مستقیم بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی وجود دارد. پس می توان گفت که بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

جدول ۴-۲۳ آزمون F (آنالیز واریانس) جهت معنادار بودن رگرسیون

متغیر	مجموع مربعات	میانگین مربعات	درجه آزادی	F	سطح معناداری
رگرسیون	۱۴,۰۹۷	۱۴,۰۹۷	۱	۲۷,۷۲۳	۰,۰۰۰
باقی مانده	۲۹,۴۹۲	۰,۵۰۸	۵۸		
کل	۴۳,۵۸۹		۵۹		

جدول ۴-۲۳ میزان متغیر مستقل یعنی دسترس پذیری در فرآیندهای بازاریابی متغیر وابسته یعنی کاربرد Big Data را نشان می دهد. همانطور که مشاهده می کنید مقدار معنی داری کوچکتر از ۰/۰۵ است بنابراین به طور کلی می توان گفت متغیر مستقل بر متغیر وابسته اثرگذار است.

جدول ۴-۲۴ ضرایب تاثیر رگرسیون

متغیر	بتا	T	سطح معناداری
مقدار ثابت	-	۲,۹۱۵	۰,۰۰۴
دسترس پذیری	۰,۵۶۹	۵,۲۶۵	۰,۰۰۰

مقدار ثابت همان عرض از مبدا است و میزان متغیر وابسته را بدون دخالت متغیر مستقل نشان می دهد. با توجه به نتایج جدول فوق می توان گفت با ارتقا یک واحد از متغیر مستقل به میزان ضریب نوشته شده متغیر وابسته ارتقا پیدا خواهد کرد. آماره  $t$  اهمیت نسبی متغیر مستقل را نشان می دهد. با توجه به مقدار این آماره و سطح خطای کمتر از ۰,۰۵ می توان گفت متغیر مورد نظر تاثیر آماری معنی داری در تبیین تغییرات متغیر وابسته داشته است. ضرایب رگرسیون نیز تاثیر متغیر مستقل بر متغیر وابسته را تایید می کند. با توجه به ضریب دسترس پذیری، فرضیه پژوهش مورد تایید قرار می گیرد. در نتیجه می توان گفت بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی رابطه معناداری وجود دارد. یعنی فرض صفر فرضیه چهارم رد و فرض یک تایید می گردد.

**فرضیه پنجم: بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی رابطه معناداری وجود دارد.**

$H_0$ : به نظر می رسد بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی رابطه معناداری وجود ندارد.

$H_1$ : به نظر می رسد بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

**جدول ۴-۲۵ ضریب همبستگی پیرسون فرضیه پنجم**

		کاربرد Big Data		مدیریت پایگاه داده
Big Data	کاربرد	سطح پیرسون	۱	۰,۶۲۹
		سطح معناداری		۰,۰۰۰
		تعداد	۶۰	۶۰
پایگاه داده	مدیریت	سطح پیرسون	۰,۶۲۹	۱
		سطح معناداری	۰,۰۰۰	
		تعداد	۶۰	۶۰

همانطور که در جدول ۴-۲۵ مشاهده می کنید، مقدار سطح معناداری فرضیه پنجم کمتر از ۰/۰۵ است؛ بنابراین بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی رابطه معناداری وجود دارد. بنابراین می توان گفت که فرضیه پنجم تایید و فرض صفر آن رد می شود. مقدار و علامت ضریب این آزمون نیز جهت و قدرت رابطه را نشان می دهد. از آنجایی که علامت ضریب مثبت است بنابراین رابطه مستقیم بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی وجود دارد. پس می توان گفت که بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی رابطه معناداری وجود دارد.

جدول ۴-۲۶ آزمون F (آنالیز واریانس) جهت معنادار بودن رگرسیون

متغیر	مجموع مربعات	میانگین مربعات	درجه آزادی	F	سطح معناداری
رگرسیون	۱۷,۲۵۵	۱۷,۲۵۵	۱	۳۸,۰۰۵	۰,۰۰۰
باقی مانده	۲۶,۳۳۴	۰,۴۵۴	۵۸		
کل	۴۳,۵۸۹		۵۹		

جدول ۴-۲۶ میزان متغیر مستقل یعنی مدیریت پایگاه داده در فرآیندهای بازاریابی متغیر وابسته یعنی کاربرد Big Data را نشان می دهد. همانطور که مشاهده می کنید مقدار معنی داری کوچکتر از ۰/۰۵ است بنابراین به طور کلی می توان گفت متغیر مستقل بر متغیر وابسته اثرگذار است.

جدول ۴-۲۷ ضرایب تاثیر رگرسیون

متغیر	بتا	T	سطح معناداری
مقدار ثابت	-	۶,۴۰۷	۰,۰۰۰
مدیریت پایگاه داده	۰,۶۲۹	۶,۱۶۵	۰,۰۰۰

مقدار ثابت همان عرض از مبدا است و میزان متغیر وابسته را بدون دخالت متغیر مستقل نشان می دهد. با توجه به نتایج جدول فوق می توان گفت با ارتقا یک واحد از متغیر مستقل به میزان ضریب نوشته شده متغیر وابسته ارتقا پیدا خواهد کرد. آماره  $t$  اهمیت نسبی متغیر مستقل را نشان می دهد. با توجه به مقدار این آماره و سطح خطای کمتر از ۰,۰۵ می توان گفت متغیر مورد نظر تاثیر آماری معنی داری در تبیین تغییرات متغیر وابسته داشته است. ضرایب رگرسیون نیز تاثیر متغیر مستقل بر متغیر وابسته را تایید می کند. با توجه به ضریب مدیریت پایگاه داده، فرضیه پژوهش مورد تایید قرار می گیرد. در نتیجه می توان گفت بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی رابطه معناداری وجود دارد. یعنی فرض صفر فرضیه پنجم رد و فرض یک تایید می گردد.

#### ۴-۵ جمع بندی فصل

در مقدمه این فصل بیان گردید که به منظور توصیف ویژگی های نمونه، ابتدا داده های جمع آوری شده با استفاده از شاخص های آمار توصیفی خلاصه و طبقه بندی می شود، و سپس با استفاده از شاخص های آمار استنباطی به تایید یا رد فرضیات پرداخته شده است. در بخش شاخص های آماری (میان، انحراف استاندارد و میانگین) متغیرهای مستقل و وابسته گزارش شده است. در واقع این شاخص ها می تواند به درک بیش تر متغیرهای پژوهش کمک کند. همچنین در بخش آمار استنباطی پس از مشخص کردن نحوه پراکنش داده ها فرضیات را با استفاده از آزمون های آماری مناسب مورد بررسی قرار داده ایم، سپس به بررسی نرمال بودن داده ها بر اساس آزمون معتبر کلمو گروف اسمیرنوف پرداخته شده است و در نهایت به آزمون فرضیه ها پرداخته شد.



# فصل پنجم: بحث، نتیجه گیری و پیشنهادات

## ۵-۱ مقدمه

پژوهش حاضر با هدف تجزیه و تحلیل کاربرد Big Data در فرآیندهای بازاریابی انجام شد. گزارش حاضر در پنج فصل تدوین گشت. فصل اول به بیان کلیاتی در خصوص ضرورت انجام تحقیق، اهداف و سوالات تحقیق و توصیف اجمالی واژگان تخصصی اختصاص یافت. در فصل دوم مبانی نظری و پیشینه تحقیق و مطالعات مرتبط شرح داده شد. روش اجرای تحقیق و توضیحاتی در باب روش و ابزار گردآوری اطلاعات در فصل سوم ارائه شد. نتایج حاصل از تجزیه و تحلیل اطلاعات نیز در قالب جداول و نمودارهایی در فصل چهارم ارائه شد. فصل حاضر با عنوان فصل پنجم، به بیان خلاصه ای از نتایج تحقیق و نیز بحث در خصوص نتایج حاصله پرداخته و پس از یک نتیجه گیری کلی به بیان محدودیت هایی می پردازد که محقق طی اجرای این پژوهش با آنها مواجه بوده است. نهایتاً پیشنهاداتی در قالب پیشنهادات علمی و کاربردی ارائه می گردد.

## جدول ۵-۱ خلاصه نتایج

ردیف	فرضیه	نتیجه
۱	بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد.	تایید
۲	بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.	تایید
۳	بین کاربرد Big Data و تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی رابطه معناداری وجود دارد.	تایید
۴	بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی رابطه معناداری وجود دارد.	تایید
۵	بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی رابطه معناداری وجود دارد.	تایید

## ۵-۳ بحث و نتیجه گیری

هدف پژوهش حاضر بررسی تجزیه و تحلیل کاربرد Big Data در فرآیندهای بازاریابی می باشد، لذا بدین منظور پرسشنامه ای بین جامعه آماری توزیع گردید و با استفاده از نرم افزار SPSS مورد تجزیه و تحلیل قرار گرفت. یافته های

حاصل از پژوهش نشان داد که فرضیه اول مبنی بر بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد، به علت آنکه مقدار سطح معناداری کمتر از ۰/۰۵ است؛ بنابراین بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد. بنابراین می توان گفت که فرضیه اول پژوهش بین کاربرد Big Data و یکپارچگی در فرآیندهای بازاریابی رابطه معناداری وجود دارد. یافته های حاصل از پژوهش نشان داد که فرضیه دوم مبنی بر بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی رابطه معناداری وجود دارد، به علت آنکه مقدار سطح معناداری کمتر از ۰/۰۵ است؛ بنابراین بین کاربرد Big Data و جذب مشتری در فرآیندهای بازاریابی رابطه معناداری وجود دارد. بنابراین می توان گفت که فرضیه دوم مورد تایید است. یافته های حاصل از پژوهش نشان داد که فرضیه سوم مبنی بر بین کاربرد Big Data و تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی رابطه معناداری وجود دارد، به علت آنکه مقدار سطح معناداری کمتر از ۰/۰۵ است؛ بنابراین بین کاربرد Big Data و تامین نیازها و خواسته های مشتریان در فرآیندهای بازاریابی رابطه معناداری وجود دارد. بنابراین می توان گفت که فرضیه سوم مورد تایید است. یافته های حاصل از پژوهش نشان داد که فرضیه چهارم مبنی بر بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی رابطه معناداری وجود دارد، به علت آنکه مقدار سطح معناداری کمتر از ۰/۰۵ است؛ بنابراین بین کاربرد Big Data و دسترس پذیری در فرآیندهای بازاریابی رابطه معناداری وجود دارد. بنابراین می توان گفت که فرضیه چهارم مورد تایید است. یافته های حاصل از پژوهش نشان داد که فرضیه پنجم مبنی بر بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی رابطه معناداری وجود دارد، به علت آنکه مقدار سطح معناداری کمتر از ۰/۰۵ است؛ بنابراین بین کاربرد Big Data و مدیریت پایگاه داده در فرآیندهای بازاریابی رابطه معناداری وجود دارد. بنابراین می توان گفت که فرضیه پنجم مورد تایید است. نتایج حاصل از مطالعات و تحقیقات مشابه بدین قرار است:

دانوس و همکاران در پژوهشی به این نتیجه رسیدند که مقدار داده های تولیدی و مرتبط در اینترنت بطور قابل توجهی رو به افزایش بوده و این پدیده باعث ایجاد چالش هایی برای آن دسته از سازمان ها شده که دوست دارند مزایای تحلیل این جریان عظیم از داده ها را بدست آورند. این موضوع بدان علت است که داده های عظیم می توانند دیدگاه های منحصر به فردی را در انواع گرایشات به سمت بازار، الگوهای خرید مشتری و همچنین روش های پایین آوردن هزینه ها و بالتبع قیمت ها را فراهم سازند و باعث تصمیمات تجاری هدفمندتری نیز شوند.

بندریان (۱۳۸۸) در پژوهشی با عنوان بازاریابی و تجاری سازی فناوری های جدید: مراحل، عوامل تسهیل کننده و کلیدی موفقیت بیان می کنند که یکی از پیچیده ترین مراحل نوآوری، مرحله انتقال یافته های تحقیقاتی به عرصه بازار و تولید است که با عنوان «انتقال فناوری از تحقیقات به تولید» مطرح می شود. واقعیت این است که هیچ الگوی قطعی و بلامنازعی در این زمینه وجود ندارد. فرایند تجاری سازی فناوری های جدید فرایندی پیچیده، نامنظم و تصادفی است و این

فرایند برای هر فناوری جدید دارای مشکلات و ویژگیهای خاص خود می باشد. تجاری سازی موفق فناوری حداقل نیازمند به تقاضای کافی، برتری بالقوه فناوری برای پاسخگویی به آن تقاضا و یک سازمان و یا فرد کارآفرین با منابع و توانمندیهای مدیریتی و بازاریابی مناسب برای تحویل محصول نهایی به بازار می باشد. یکی از مهمترین ارکان فرایند تجاری سازی فناوری، بازاریابی فناوری می باشد. بازاریابی فناوری از نوع بازاریابی تعاملی است، بدین معنی که کیفیت فناوری تا حد زیادی به کیفیت تعامل فروشنده و خریدار بستگی دارد.

## ۴-۵ محدودیت های پژوهش

هر تحقیق و پژوهشی در بطن خود یکسری محدودیت ها خواهد داشت که در مسیر تحقق آن ظاهر شده و امر پژوهش را با مشکلاتی روبرو می نماید که تعمیم نتایج را با دشواری همراه می سازد. پژوهش حاضر نیز از این قاعده مستثنی نبوده و دارای محدودیت هایی به شرح زیر بوده است:

- (۱) عدم همکاری برخی افراد جامعه آماری جهت تکمیل پرسشنامه
- (۲) برداشت های متفاوت از گویه های پرسشنامه ها
- (۳) نتایج به قلمرو زمانی و مکانی دیگر قابل تغییر نیست.
- (۴) محدودیت در عدم کنترل برخی از متغیرهای تاثیرگذار بر نتایج پژوهش
- (۵) محدود نمودن جامعه آماری

## ۵-۵ پیشنهادات

نتایج حاصل از هر پژوهشی به امید ادامه یافتن راه تحقیق و پژوهش در خصوص آن موضوع و بهره برداری از نتایج آن به جامعه پژوهشگران و مسئولین ذی صلاح آن موضوع ارائه می گردد. از اینرو ارائه هر نوع پیشنهادی در این گزارشات می تواند راه را برای مطالعات بعدی و نیز تصمیم گیری های اجرایی در آن خصوص هموار سازد. در این پژوهش نیز پیشنهاداتی در قالب پیشنهادات پژوهشی جهت کارهای مطالعاتی بعدی و نیز پیشنهادات کاربردی برای تصمیم گیری های اجرایی و توجه مسئولین ذی ربط به آن حوزه ارائه شده است.

## ۵-۵-۱ پیشنهادات کاربردی

بر اساس یافته‌های پژوهش حاضر، در این قسمت سعی شده است تا به بیان پیشنهاداتی کاربردی پرداخته شود. امید است دستگاه‌های مسئول و سازمان‌های ذی‌ربط به این پیشنهادات به دیده دقت بنگرند.

- ✓ استفاده از مدل‌های مناسب مدیریت پایگاه داده در فرآیندهای بازاریابی به منظور بهبود فرآیند کسب و کار
- ✓ آشنا نمودن مدیران و کارشناسان نسبت به مزایای ناشی از کاربرد Big Data در فرآیندهای بازاریابی از طریق برگزاری سمینار و همایش‌ها
- ✓ بالا بردن توانایی مدیران در بکارگیری سیستم Big Data به منظور بهبود فرآیندهای بازاریابی
- ✓ استفاده از تکنیک‌های پیشرفته ارزیابی و تحلیل داده‌ها جهت بهبود فرآیندهای بازاریابی
- ✓ تخصیص بودجه از سوی دولت و ارگان‌های مربوطه به منظور استفاده Big Data در شرکت‌های کسب و کار به منظور بهبود فرآیند بازاریابی

## ۵-۵-۲ پیشنهادات آتی

- ✓ بررسی مزایا و معایب ناشی از بکارگیری Big Data در طی پژوهش‌های آتی
- ✓ بررسی موانع ناشی از بکارگیری Big Data در طی پژوهش‌های آتی

## منابع فارسی :

۱. 'استراس، آنسلم، جولیت کوربین. اصول روش تحقیق کیفی. ترجمه بیوک محمدی. تهران پژوهشگاه علوم انسانی و مطالعات فرهنگی. ۱۳۹۰. ص ۱۴۴
۲. آذریان، الهام سادات . ۱۳۹۶. چالش امنیت کلان داده در فناوری ابر . اولین همایش ملی مرکز داده زیرساخت و سرویس ها
۳. پرهیزگار، محمد مهدی، آقاجانی افروزی، علی اکبر. روش شناسی تحقیق پیشرفته در مدیریت با رویکرد کاربردی، دانشگاه پیام نور. ۱۳۹۰ ص ۹۲.
۴. حافظ نیا، محمد رضا. مقدمه ای بر روش تحقیق در علوم انسانی، ناشر سازمان مطالعه و تدوین کتب علوم انسانی دانشگاه ها(سمت). ۱۳۸۹. ص ۹۳
۵. خالد علی دوست. ۱۳۹۴. اهمیت کلان داده و کاربردهای آن در حوزه های مختلف .چهارمین همایش پژوهش های نوین در علوم و فناوری.
۶. خدیور، آمنه. امیریان، بهاره. ۱۳۹۵. ارزیابی بلوغ کلان داده در صنعت بیمه کشور. دولتی - وزارت علوم، تحقیقات، و فناوری - دانشگاه الزهرا (س) - دانشکده اقتصاد و حسابداری - کارشناسی ارشد
۷. دلاور، علی. روش تحقیق در روان شناسی و علوم تربیتی، ناشر ویرایش. ۱۳۸۶. ص ۶۵
۸. رضاپورنیاری، مریم. ۱۳۹۶. کاربرد کلان داده در اینترنت اشیا و تأثیر آن بر صنعت ورودکست. رسانه هوشمند، رسانه رایانه.
۹. سرمد، زهره، بازرگان، عباس، حجازی، الهه. روش های تحقیق در علوم رفتاری. تهران. انتشارات آگه. ۱۳۹۴ ص ۷۷
۱۰. عنایتی، الهام. ۱۳۹۳. کاربرد رایانش ابری در داده های حجیم . بیستمین کنفرانس ملی سالانه انجمن کامپیوتر ایران.
۱۱. فرشاد نوریان، سمانه حجازی . ۱۳۹۲. کاربرد کلان-داده ها در نقد توسعه مبتنی بر حمل و نقل عمومی .مطالعات شهری. ۸۳-۹۱. ۸.
۱۲. محمودزاده، ابراهیم. صحرایی ، مهدی . قوچانی خراسانی، محمدمهدی. ۱۳۹۶. تدوین استراتژی کلان داده در تحلیل شبکه های اجتماعی برای پیش بینی بحران .مدیریت بحران. ۷۷-۹۱.

۱۳. میثم وکیلی، محمدجواد کارگر. ۱۳۹۳. بهبود توازن بار در مدل نگاشت کاهش جهت پردازش کلان داده در خوشه های ناهمگن. دومین همایش ملی پژوهش های کاربردی در علوم کامپیوتر و فناوری اطلاعات

۱۴. میرزایی، حمیدرضا. روش تحقیق، انتشارات شایسته گستر. ۱۳۸۸، ص ۷۴  
۱۵. هلیلی، خداداد. موسوی، فاطمه سادات. سلطان پور، محمدرضا. ۱۳۹۵. لزوم بکارگیری فناوری کلان داده در سامانه های C4I و بررسی چالش های آن. نهمین کنفرانس ملی فرماندهی و کنترل ایران

#### منابع لاتین

۱۶. "A Guide to the Business Analysis Body of Knowledge (BABOK Guide) Version 2.0", IIBA, 2009
۱۷. Alexandors Labrinidis and HV Jagadish. Challenges and opportunities with big data. Proceeding of the VLDB Endowment, 5(12):2032-2۰۳۳, ۲۰۱۲.
۱۸. Big data. <http://www.nature.com/news/specials/bigdata/index.html>, 2008.
۱۹. Barbara Carkenord, 2010.. "Why Does a Project Need a Project Manager and a Business Analyst",
۲۰. D Agrawal, P Bernstein, E Bertino, S Davidson, U Dayal, M Franklin, J Gehrke, L Haas, A Halevy, J Han, et al. Challenges and opportunities with big data. A community white paper developed by leading researchers across the united states, 2012.
۲۱. Douglas Laney. 3-d data management: Controlling data volume, velocity and variety. META Group Research Note, February, 6, 2001.
۲۲. Drowning in numbers- digital data will flood the planet- and help us understand it better. bigdata-0, 20۱۱.
۲۳. D Boyd, K Crawford 2012.. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. Journal Information, Communication & Society Volume 15, Issue 5

٢٤. Erik Meijer. The world according to linq. Communications of the ACM, ٥٤(١٠):٤٥-٥١, ٢٠١١.
٢٥. Fan J, Han F, Liu H. Challenges of big data analysis. National science review. 2014 Jun 1;1(2):p 293
٢٦. Fact sheet: Big data across the federal government.  
[http://www.whitehouse.gov/sites/default/files/microsites/ostp/big\\_data\\_fact\\_sheet\\_3\\_29\\_2012.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_fact_sheet_3_29_2012.pdf), 2012.
٢٧. Felix Castro, Alfredo Vellido, Angela Nebot, and Francisco Mugica, “Applying Data Mining Techniques to e-Learning Problems,” Studies in Computational Intelligence, Volume 62, 2007, pp. 183-2٢١
٢٨. Joseph Beck and Jack Mostow, “How Who Should Practice: Using Learning Decomposition to Evaluate the Efficacy of Different Types of Practice for Different Types of Students,” Proceedings of the 9th International Conference on Intelligent Tutoring Systems, 2008, pp 353-3٦٢.
٢٩. James Manyika, McKinsey Global Institute, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, 2011.
٣٠. James Manyika, McKinsey Global Institute, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, 2011.
٣١. James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Byers, “Big Data: The Next Frontier for Innovation, Competition, and Productivity,” McKinsey Global Institute, May, 2011



۳۲. Jeremy Ginsberg, Matthew H Mohebbi, Rajan S Patel, Lynnette Brammer, Mark S Smolinski, and Larry Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457(7232):1012-1014, ۲۰۰۸.
۳۳. John Gantz and David Reinsel. Extracting value from chaos. IDC iView, pages 1-12, 2011.
۳۴. John H Howard, Michael L Kazar, Sherri G Menees, David A Nichols, Mahadev Satyanarayanan, Robert N Sidebotham, and Michael J West. Scale and performance in a distributed file system. *ACM Transactions on Computer Systems (TOCS)*, 6(1):51-81, ۱۹۸۸.
۳۵. Kenneth Cukier. Data, data everywhere: A special report on managing information. *Conomist Newspaper*, 2010.
۳۶. Labrinidis A, Jagadish HV. Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*. 2012 Aug ;p 67
۳۷. Lings N. I., 2004, Internal market orientation-Construct and consequences, *Journal of Business Research*, p 57
۳۸. Liran Einav and Jonathan Levin, "The Data Revolution and Economic Analysis," Working Paper, No. 19035,
۳۹. M Grobelnik. Big data tutorial. <http://videolectures.net/eswc2012grobelnikbigdata/>, 2012.
۴۰. Mark beyer. Garther says solving 'big data' challenge involves more than just managing volumes of data. *Gartner*. <http://www.gartner.com/it/page.jsp>, 2011.
۴۱. Michael Fitzgerald, "Tech Hotshots: The rise of IT Business Analyst", [www.computerworld.com](http://www.computerworld.com), July 2012.
۴۲. Noguchi Yuki. Following digital breadcrumbs to big data gold. *thedigitalbreadcrumbs-that-lead-to-big-data*, 2011.
۴۳. National Bureau of Economic Research, 2013, 19035;

٤٤. Noguchi Yuki. The search for analysts to make sense of big data.  
<http://www.npr.org/2011/11/30/142893065/the-searchforanalysts-to-make-sense-of-big-data>,2011.
٤٥. O.R Team. Big Data Now: Current Perspectives from O'Reilly Radar.O'Reilly Media,2011.
٤٦. Paul Zikopoulos, Chris Eaton, et al. Understanding big data: Analytics for enterpris class hadoop and streaming data. McGraw-Hill Osborne Media, ٢٠١١.
٤٧. PenelopePeterson,andEvaBaker,eds.,InternationalEncyclopediaoEducation, ThirdEdition,Oxford,United Kingdom:Elsevier, 2012.
٤٨. Richard Kemp. Legal Aspects of Managing Big Data . September .2014.
٤٩. Rick Cattell. Scalable sql and nosql data store. ACM SIGMOD Record.39(4):12-27,2011.
٥٠. Ryan Baker, "Data Mining for Education," Barry McGaw,
٥١. Roger S. Pressman,2012. "Software Engineering, A practitioner's Approach", McGraw Hill, Fifth Edition.
٥٢. Sharona Levy and Iri Wilensky, "Mining Students' Inquiry Actions for Understanding of Complex Systems," Computers &Education, Volume 56, ٢٠١١, pp. ٥٥٦-٥٧٣.
٥٣. Steve Lohr. The age of big data. New York Times, 11, 2012.
٥٤. Special online collection: Dealing with big data.  
<http://www.sciencemag.org/sit/special/data/>,2011.
٥٥. Sweeney J. C. and Soutar G., 2001, Consumer perceived value: the development of multiple item scale, Journal of Retailing, p 77
٥٦. S Erevelles, N Fukawa, L Swayne - Journal of Business Research, 2016 Big Data consumer analytics and the transformation of marketing .Journal of Business Research. Volume 69, Issue 2, Pages 897-9٠٤

- ๑๗. U.S. Department of Education Office of Educational Technology,  
“Enhancing Teaching and Learning Through Educational Data Mining and  
Learning Analytics,” 2012.
- ๑๘. Viktor Mayer- Schonberger and Kenneth Cukier, Big Data: A Revolution  
That Will Transform How We Live, Work, and Think, (Houghton Mifflin  
Harcourt, 2013)
- ๑๙. Viktor Mayer-Schonberger and Kenneth Cukier. Big Data: A Revolution  
that will Transform how We live, and Think. Eamon Dolan/Houghton  
Mifflin Harcourt, 2013.

به نام خدا

فرم نظر سنجی پایان نامه کارشناسی ارشد

" تجزیه و تحلیل کاربرد Big Data در فرآیند های بازاریابی "

با سلام و احترام؛ خواهشمندیم سوال های زیر را به دقت خوانده و مناسب ترین پاسخ را در محل مربوطه علامت بزنید. بی تردید همکاری و پاسخ های دقیق شما ما را در رسیدن به اهداف و انجام وظایف مان یاری خواهد رساند اطلاعات شخصی پاسخگویان تنها در جهت پربار سازی پژوهش است و این اطلاعات نزد پژوهشگر به صورت محرمانه و محفوظ خواهد ماند.

پرسشنامه اطلاعات فردی				
جنسیت	<input type="checkbox"/> زن	<input type="checkbox"/> مرد		
سن	<input type="checkbox"/> کمتر از ۲۵ سال	<input type="checkbox"/> ۲۵ تا ۳۰	<input type="checkbox"/> ۳۰ تا ۳۵	<input type="checkbox"/> بالاتر از ۴۰
تحصیلات	<input type="checkbox"/> کاردانی	<input type="checkbox"/> کارشناسی	<input type="checkbox"/> کارشناسی ارشد	<input type="checkbox"/> دکتری
وضعیت تاهل	<input type="checkbox"/> مجرد	<input type="checkbox"/> متاهل		
سابقه خدمت	<input type="checkbox"/> کمتر از ۵ سال	<input type="checkbox"/> ۵ تا ۱۰	<input type="checkbox"/> ۱۰ تا ۱۵	<input type="checkbox"/> ۱۵ تا ۲۰
	<input type="checkbox"/> بالاتر از ۲۰ سال			

خیلی کم	کم	تا حدودی	زیاد	خیلی زیاد	سوالات
					وجود زیر ساخت مناسب به منظور تحلیل Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
					کسب درآمد از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
					افزایش رضایت مشتری از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
					استفاده از روش های مناسب برای مدیریت داده ها در جهت بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
					توسعه برنامه های کاربردی از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
					کسب بینش غنی تر و عمیق تر و موفقیت در رقابت از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
					استفاده از تکنیک های پیشرفته ارزیابی و تحلیل داده ها از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
					آسان کردن فرآیند تصمیم سازی از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
					درک جامعی از نیازمندی های مشتریان در جهت بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
					اتخاذ تصمیمات هوشمندانه و همسو با نیازهای مشتریان از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
					روی آوردن شرکت ها و افراد به شکل های جدیدی از زندگی دیجیتالی در جهت بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
					تعیین راهبرد های مناسب برای بهره برداری صحیح از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟

				بهبود ساختارها و سودآوری شرکت های کسب و کار از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
				افزایش سرعت تولید داده های ورودی و خروجی از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
				بهینه سازی فرآیندهای کسب و کار از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
				تجاری سازی فناوری های جدید از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
				دسترسی به اطلاعات مناسب و سودمند از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
				کارآمدتر کردن فرآیندهای کسب و کار از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
				درک خواسته های مشتریان و هدف قرار دادن نیازهای آنان از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
				افزایش محدوده تنوع و منابع داده ها از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
				کنترل و پردازش داده ها از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
				صرفه جویی در وقت و هزینه از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
				استفاده از تکنیک های پیشرفته ارزیابی و تحلیل داده ها از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
				انجام سریع اقدامات لازم جهت عقب نماندن از فرصت های بازار از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟
				جلوگیری از خطاهای انسانی از طریق بکارگیری Big Data تا چه حد در فرآیند بازاریابی تاثیرگذار می باشد؟



**University of Tehran**

**Irrigation and Reclamation Engineering Department**

By

**Shabnam jalali**

Under Supervision of **Dr. Aslani**

Adviser: **Dr. Nazari**

**A thesis submitted to the Graduate Studies Office in partial fulfillment of the requirements  
for the degree of Master of Science (Doctor of Philosophy) in**

**Big data usage analysis in marketing processes**

**January, 2018**

## Abstract

It is absolutely important to create necessary capabilities to maintain work environment and competition with other competitors in nowadays complex world which is continuously changing in various technological fields.

Concerning the conditions on today's markets, it is noticeable for us to identify customers' ever changing necessities which make the most important enterprises of active companies. Since each of these companies can survive only if they can coordinate with demands and compete with others. By using analysis of more capacities of data, we can have better and more developed analysis for different goals, such as trade, medical and security goals and can get proper results.

The aim of this research is to analyze Big Data usage on marketing and competition processors. To do this first, we will identify the major factors affecting on organization business from competition and marketing point of view.

Then, we'll consider to identify activity mechanism and the frames of Big Data usage. The present research is considered a descriptive one by expressing unique features of Big Data tools. This leads to this point that the result can be used in industries and economical activities, so it is applicable.

A research questionnaire was distributed among statistical population and was analyzed by SPSS software.

The finding showed that there is a meaningful relationship between using Big Data consistency in marketing processors. Since there is a lower meaningful level less than %5, so there exists a meaningful relationship between using Big Data and consistency in marketing processors.



Key words: Big Data – technology – marketing processors.