

فهرست مطالب

بخش اول.....	۲۱
شروع کار با علم داده	۲۱
فصل اول.....	۲۳
نگاهی اجمالی بر علم داده‌ها	۲۳
بررسی راهکارهای پیشنهادی علم داده	۲۴
ایجاد یک تیم درون سازمانی.....	۲۴
برون‌سپاری کارها به مشاورین خصوصی علم داده	۲۵
استفاده از راه‌حل‌های ابر محور.....	۲۶
مزایای محسوس علم داده	۲۶
فصل دوم.....	۲۷
بررسی زیرساخت‌های مهندسی داده	۲۷
تعریف کلان داده با ۴ ویژگی اصلی آن	۲۸
تنوع داده‌ها	۳۰
منابع کلان داده کدامند؟	۳۱
تفاوت بین مهندسی داده و علم داده	۳۱
تعریف علم داده	۳۲
تعریف مهندسی داده	۳۲
مقایسه‌ی متخصصین داده و مهندسان داده	۳۳
محدود کردن داده‌ها با استفاده از <i>hadoop</i> و <i>mapreduce</i>	۳۴
نگاهی عمیق‌تر به <i>mapreduce</i>	۳۴

- نگاشت داده‌ها ۳۵
- کاهش داده ۳۵
- Hadoop چیست؟ ۳۶
- معرفی چارچوب‌های پردازش بلادرنگ ۳۹
- معرفی پلتفرم‌های پردازش موازی عظیم Mpp ۴۰
- معرفی پایگاه داده‌های NoSQL ۴۱
- بررسی یک نمونه عملی از مهندسی داده ۴۲
- شناسایی چالش‌های مربوط به کسب‌وکار در این زمینه ۴۲
- حل مشکلات مربوط به کسب‌وکار با استفاده از مهندسی داده ۴۳
- چه مزایایی برای این شرکت حاصل شد؟ ۴۳
- فصل سوم ۴۵
- کاربرد علم داده در صنعت و کسب‌وکار ۴۵
- ترکیب بینش‌های حاصل‌شده از داده‌ها در فرآیندهای کسب‌وکار ۴۶
- مزایای علم داده‌ی کسب‌وکار محور ۴۶
- شناسایی چالش‌های رایج در تجزیه و تحلیل ۴۷
- گردآوری داده‌های خام و تبدیل آن‌ها به بینش عملی ۴۷
- اقدام بر اساس بینش‌های حاصل‌شده ۴۹
- تفاوت هوش تجاری و علم داده ۵۰
- تعریف هوش تجاری ۵۱
- نگاهی به انواع داده‌ی مورد استفاده در هوش تجاری ۵۶
- بررسی تکنولوژی و مهارت‌هایی که در هوش تجاری مفید هستند ۵۶
- فرایند ETL ۵۸
- تعریف علم داده کسب‌وکار محور ۵۸
- بررسی انواع داده‌های مفید در علم داده‌ی کسب‌وکار محور ۵۹

- چه فناوری و مهارت‌هایی در علم داده کسب‌وکار محور مفید هستند؟ ۶۰
- تفاوت‌های اصلی هوش تجاری و علم داده کسب‌وکار محور ۶۱
- علم داده در تجارت (داستان موفقیت یک تجارت مبتنی بر داده) ۶۳
- بخش دوم** ۶۵
- استفاده از علم داده در راستای استخراج بینش از داده‌ها ۶۵
- فصل چهارم** ۶۷
- معرفی آمار و احتمالات ۶۷
- معرفی مفاهیم اساسی احتمال ۶۸
- بررسی رابطه بین احتمال و آمار استنباطی ۶۸
- آشنایی با برخی توزیع‌های احتمال رایج ۷۰
- معرفی رگرسیون خطی ۷۲
- آشنایی با مدل‌های ساده رگرسیون خطی ۷۲
- آموزش ساخت خط رگرسیون با اندازه متناسب ۷۴
- روش‌هایی ساده جهت محاسبه رگرسیون حداقل مربعات معمولی شبیه‌سازی ۷۵
- چگونه این عدد تصادفی به رنگ چشم مرتبط است؟ ۷۹
- استفاده از شبیه‌سازی برای ارزیابی خواص یک آماره آزمون ۸۱
- استفاده از شبیه‌سازی مونت‌کارلو برای ارزیابی خواص یک برآوردگر ۸۲
- معرفی تجزیه و تحلیل سری‌های زمانی ۸۵
- درک الگوهای سری زمانی ۸۵
- مدل‌سازی داده‌های سری زمانی تک متغیره ۸۶
- فصل پنجم** ۸۹
- خوشه‌بندی و طبقه‌بندی** ۸۹
- معرفی اصول خوشه‌بندی و طبقه‌بندی ۹۰
- درک الگوریتم‌های خوشه‌بندی ۹۱

۹۶.....	<i>k-means</i> خوشه‌بندی با استفاده از الگوریتم
۹۷.....	روش <i>KDE</i>
۹۸.....	خوشه‌بندی با الگوریتم‌های سلسله مراتبی و محلی
۱۰۰.....	طبقه‌بندی داده‌ها با درخت تصمیم و الگوریتم‌های جنگل تصادفی
۱۰۳.....	فصل ششم
۱۰۳.....	خوشه‌بندی و طبقه‌بندی
۱۰۳.....	با استفاده از الگوریتم نزدیک‌ترین همسایه
۱۰۴.....	مفهوم بخشی به داده‌ها با استفاده از تجزیه و تحلیل نزدیک‌ترین همسایه
۱۰۵.....	درک اهمیت خوشه‌بندی و طبقه‌بندی
۱۰۶.....	طبقه‌بندی داده‌ها با استفاده از الگوریتم‌های میانگین نزدیک‌ترین همسایه
۱۰۷.....	مقایسه شباهت‌های میانگین با تحلیلگر کسب و کار <i>Stu</i>
۱۰۹.....	طبقه‌بندی با الگوریتم‌های <i>k</i> آمین-نزدیک‌ترین همسایه
۱۱۱.....	چه موقع باید از الگوریتم <i>KNN</i> استفاده کنیم؟
۱۱۲.....	کاربردهای رایج الگوریتم <i>KNN</i>
۱۱۲.....	استفاده از نزدیک‌ترین فاصله برای استخراج مفاهیم الگوهای نقاط
۱۱۳.....	حل مسائل دنیای واقعی به وسیله الگوریتم‌های نزدیک‌ترین همسایه
۱۱۶.....	استفاده از روش‌های عددی در علم داده‌ها
۱۱۷.....	بسط چندجمله‌ای تیلور
۱۱۹.....	دوبخشی کردن توابع با استفاده از الگوریتم جستجوی دوبخشی
۱۲۰.....	مدل‌سازی ریاضیاتی با زنجیره‌های مارکوف و روش‌های تصادفی
۱۲۳.....	فصل هفتم
۱۲۳.....	مدل‌سازی داده‌های فضایی با آمار
۱۲۴.....	تولید سطح قابل پیش‌بینی از نقاط داده‌های فضایی
۱۲۵.....	آشنایی با (X, Y, Z) مدل داده‌های فضایی

۱۲۶ معرفی <i>kriging</i>
۱۲۷ کردن به منظور درون‌یابی <i>kriging</i> خودکار
۱۲۷ انتخاب و استفاده از مدل‌های مناسب برای درون‌یابی <i>kriging</i> صریح و تعریف‌شده
۱۲۷ میانگیری با مدل‌های <i>variogram</i> صریح و تعریف‌شده
۱۲۸ تخمین <i>variogram</i>
۱۲۹ نگاهی دقیق به روش <i>kriging rabbit hole</i>
۱۲۹ بررسی روش‌های تخمین سطح <i>kriging</i> های معمولی
۱۳۰ بررسی روش‌های تخمین سطح رگرسیون <i>kriging</i>
۱۳۰ بررسی تخمین زن‌های سطح <i>kriging</i> بلوکی
۱۳۲ بررسی تخمین زن‌های سطح <i>cokriging</i>
۱۳۳ انتخاب بهترین روش تخمین در <i>kriging</i>
۱۳۵ تحلیل باقی‌مانده‌ها برای تعیین بهترین مدل مناسب
۱۳۸ شناخت گزینه‌های موجود در <i>kriging</i>
۱۳۸ استفاده از تحلیل سطح روند بر داده‌های فضایی
۱۴۱ بخش سوم
۱۴۱ ایجاد بصری سازی‌های معنادار از داده‌ها
۱۴۳ فصل هشتم
۱۴۳ به‌کارگیری اصول طراحی بصری‌سازی داده
۱۴۴ شناخت انواع بصری‌سازی
۱۴۴ توصیف داستانی داده‌ها برای تصمیم‌گیرندگان سازمانی:
۱۴۵ نمایش داده‌ها برای تحلیل‌گران
۱۴۵ طراحی تابلوی نمایش داده برای فعالان این زمینه
۱۴۵ تمرکز بر مخاطبان
۱۴۸ انتخاب مناسب‌ترین حالت طراحی

۱۴۸ استفاده از بصریسازی برای درک دقیق محاسباتی
۱۴۹ استفاده از طراحی برای ایجاد شور و اشتیاق در مخاطب
۱۵۱ استفاده از داده‌ها برای ایجاد متن
۱۵۱ ایجاد متن با استفاده از حاشیه‌نویسی
۱۵۳ ایجاد متن با استفاده از عناصر گرافیکی
۱۵۳ چه زمانی باید از حالت متقاعدکننده استفاده کنیم؟
۱۵۳ انتخاب مناسب‌ترین نوع داده گرافیکی
۱۵۴ بررسی نمودار استاندارد گرافیکی
۱۵۷ بررسی گرافیک مقایسه‌ای
۱۶۱ بررسی طرح‌های آماری
۱۶۳ بررسی ساختارهای توپولوژی
۱۶۵ بررسی طرح‌ها و نقشه‌های فضایی
۱۶۷ انتخاب گرافیک داده شما
۱۶۸ تعیین دامنه سؤالات
۱۶۸ در نظر گرفتن رسانه و کاربر
۱۶۸ نگاهی مجدد به گام نهایی بیندازید
۱۶۹ فصل نهم
۱۶۹ استفاده از <i>DOM.js</i> برای بصریسازی داده
۱۷۰ معرفی کتابخانه <i>DOM.js</i>
۱۷۱ چه زمانی باید از <i>DOM.js</i> استفاده کنیم؟
۱۷۲ شروع به کار در <i>DOM.js</i>
۱۷۲ بررسی <i>HTML</i> و <i>DOM</i>
۱۷۳ بررسی <i>JavaScript</i> و <i>SVG</i>
۱۷۵ بررسی سرورهای وب و <i>PHP</i>

۱۷۶	شناخت مفاهیم پیشرفته‌تر و شیوه عملکرد <i>D3.js</i>
۱۸۰	بررسی ساختار زنجیره‌ای
۱۸۱	بررسی مقیاس‌ها
۱۸۲	مروری بر انتقال‌ها و تعاملات
۱۸۵	فصل دهم
۱۸۵	اپلیکیشن مبتنی بر وب برای طراحی بصری سازی
۱۸۷	کار با تحلیلگر واتسون <i>IBM</i>
۱۸۸	بصری سازی و کار تیمی با استفاده از <i>PLOTLY</i>
۱۹۱	بصری سازی داده‌های فضایی با استفاده از ابزارهای جغرافیایی آنلاین
۱۹۲	ایجاد نقشه‌های خیره‌کننده با <i>Open Heat Map</i>
۱۹۳	ایجاد نقشه و تحلیل داده‌های فضایی با <i>CartoDB</i>
۱۹۵	بصری سازی با ابزار متن‌باز: پلتفرم‌های بصری سازی داده مبتنی بر وب
۱۹۵	ایجاد داده‌های گرافیکی زیبا با جدول‌های <i>Google fusion</i>
۱۹۶	استفاده از <i>iCharts</i> برای بصری‌سازی داده مبتنی بر وب
۱۹۷	استفاده از <i>RAW</i> برای بصری‌سازی داده مبتنی بر وب
۱۹۹	چه زمانی باید از اینفوگرافیک استفاده کنیم؟
۱۹۹	ساختن اینفوگرافیک‌های جذاب با استفاده از <i>Infogr.am</i>
۲۰۱	مهاجرت درون ایالتی
۲۰۱	ایجاد اینفوگرافیک‌های جذاب با استفاده از <i>PiktoChart</i>
۲۰۳	فصل یازدهم
۲۰۳	بررسی بهترین روش‌ها در طراحی داشبورد
۲۰۴	تمرکز بر مخاطبین
۲۰۵	طرح کلی
۲۰۶	طراحی جزئیات

۲۰۹	فصل دوازدهم
۲۰۹	ساخت نقشه با استفاده از داده‌های فضایی
۲۱۰	آشنایی با اساس GIS
۲۱۲	فرمت‌های فایل‌های GIS
۲۱۵	درک مفهوم نقشه و سیستم‌های مختصات
۲۱۸	بافر و توابع مجاورت
۲۱۹	استفاده از تحلیل پوشش لایه‌ای
۲۲۰	طبقه‌بندی مجدد داده‌های فضایی
۲۲۱	عملیات ساده اعمال‌شده بر ویژگی‌های دارای هم‌پوشانی
۲۲۱	کار با QGIS منبع باز
۲۲۱	آشنایی با رابط QGIS
۲۲۲	اضافه کردن یک لایه‌برداری در QGIS
۲۲۳	نمایش داده در QGIS
۲۲۹	بخش چهارم
۲۲۹	محاسبات در علم داده
۲۳۱	فصل سیزدهم
۲۳۱	استفاده از PYTHON برای علم داده
۲۳۲	آشنایی با مفاهیم اساسی در PYTHON
۲۳۳	معرفی انواع داده PYTHON
۲۳۴	اعداد در PYTHON
۲۳۴	رشته‌ها در PYTHON
۲۳۴	لیست‌ها در PYTHON
۲۳۵	تاپل‌ها در PYTHON
۲۳۵	مجموعه‌ها در PYTHON

۲۳۶	واژه‌نامه‌ها در <i>PYTHON</i>
۲۳۶	استفاده حلقه‌ها در <i>PYTHON</i>
۲۳۷	آشنایی با توابع و کلاس‌ها
۲۳۷	معرفی توابع
۲۳۹	مزیت‌های استفاده از کلاس‌ها
۲۴۱	معرفی کتابخانه‌ی <i>Numpy</i>
۲۴۳	آشنایی بیشتر با کتابخانه <i>SciPy</i>
۲۴۴	استفاده از <i>MatPlotLib</i> برای بصری‌سازی داده
۲۴۵	استفاده از <i>PYTHON</i> برای تحلیل داده - یک نمونه عملی
۲۴۶	نصب <i>PYTHON</i> بر روی سیستم‌عامل‌های <i>Windows</i> و <i>Mac</i>
۲۴۷	بارگذاری فایل‌های <i>CSV</i>
۲۴۸	محاسبه میانگین وزنی
۲۵۰	رسم خطوط روند
۲۵۳	فصل چهاردهم
۲۵۳	استفاده از زبان <i>R</i> در علم داده
۲۵۴	تسلط بر لغات مربوط به زبان <i>R</i>
۲۵۶	محاسبات مربوط به علم داده
۲۵۷	آشنایی بیشتر با توابع و عملگرها
۲۶۰	پیمایش در <i>R</i>
۲۶۱	آشنایی با عملکرد اشیا
۲۶۴	نگاهی بر پکیج‌های <i>R</i>
۲۶۴	نگاهی بر پکیج‌های مشهور تجزیه و تحلیل آماری
۲۶۵	بصری‌سازی، نگاشت و نمایش نموداری در <i>R</i>
۲۶۵	بصری‌سازی آماری <i>R</i> با استفاده از <i>ggplot2</i>

۲۶۶.....	تجزیه و تحلیل شبکه‌ها با استفاده از <i>igraph</i> و <i>statnet</i>
۲۶۷.....	نقشه‌برداری و تجزیه و تحلیل الگوی نقطه‌ای فضایی با به‌کارگیری <i>spatstat</i>
۲۶۹.....	فصل پانزدهم
۲۶۹.....	استفاده از <i>SQL</i> در علم داده
۲۷۰.....	آغاز کار با <i>SQL</i>
۲۷۰.....	بررسی پایگاه داده‌های رابطه‌ای و <i>SQL</i>
۲۷۴.....	آشنایی با مفهوم طراحی پایگاه داده
۲۷۴.....	تعریف انواع داده
۲۷۵.....	توجه به قواعد جامعیت
۲۷۵.....	نرمال سازی پایگاه داده
۲۷۷.....	استفاده از <i>SQL</i> و توابع آن در علم داده
۲۷۷.....	یکپارچه‌سازی <i>SQL</i> ، <i>PYTHON</i> و اکسل در استراتژی علم داده
۲۷۷.....	استفاده از توابع <i>SQL</i> در علم داده
۲۸۲.....	متن کاوی با استفاده از <i>SQL</i>
۲۸۲.....	فصل شانزدهم
۲۸۲.....	نرم افزارهای متناسب با علم داده
۲۸۴.....	مزیت‌های تسهیل کننده‌ی اکسل
۲۸۵.....	استفاده از اکسل برای درک سریع داده‌ها
۲۸۶.....	استفاده از فیلتر در اکسل
۲۸۶.....	فرمت بندی شرطی برای نشانه‌گذاری داده‌های خارج از محدوده و روندها
۲۸۷.....	رسم نمودار در اکسل برای مشاهده داده‌های خارج از محدوده و روندها
۲۸۹.....	فرمت بندی مجدد و خلاصه‌سازی با استفاده از جداول محوری
۲۹۰.....	خودکارسازی وظایف اکسل با استفاده از ماکرو (دستورات به کار رفته در <i>Excel</i>)
۲۹۲.....	استفاده از <i>KNIME</i> برای تحلیل پیشرفته داده‌ها

۲۹۷.....	بخش پنجم.....
۲۹۷.....	اعمال مهارت‌های حوزه‌های خاص.....
۲۹۷.....	برای حل مسائل حقیقی با استفاده از علم داده.....
۲۹۹.....	فصل هفدهم.....
۲۹۹.....	بررسی تأثیر علم داده در خبرنگاری.....
۳۰۱.....	چه کسی داده‌ها را تولید میکند؟.....
۳۰۳.....	نگاهی دقیق‌تر به اهمیت داستان خبری.....
۳۰۳.....	چرا موضوعات شما برای مخاطبان اهمیت دارد؟.....
۳۰۵.....	بهترین زمان را برای انتشار مطالب انتخاب کنید.....
۳۰۶.....	چگونه باید داستان‌های نهفته در داده‌های خود را دریابید؟.....
۳۰۷.....	چگونه باید داستان داده محور خود را ارائه دهید؟.....
۳۰۷.....	جمع‌آوری داده.....
۳۰۸.....	استخراج داده‌ها برای داستان‌های خبری.....
۳۰۸.....	بررسی روندهای غیر معمول و داده‌های خارج از محدوده.....
۳۱۱.....	تأکید کردن بر داستان از طریق بصرسازی.....
۳۱۵.....	فصل هجدهم.....
۳۱۵.....	نگاهی به علم داده محیطی.....
۳۱۶.....	مدل‌سازی تعامل انسان و محیط با استفاده از هوش محیطی.....
۳۱۶.....	نگاهی به انواع مشکلات حل شده.....
۳۱۷.....	تعریف هوش محیطی.....
۳۱۸.....	شناسایی سازمان‌های بزرگ که در حوزه هوش محیطی کار میکنند.....
۳۱۹.....	تأثیرگذاری از طریق هوش محیطی.....
۳۲۱.....	مدل‌سازی منابع طبیعی خام.....
۳۲۱.....	بررسی مدل‌سازی منابع طبیعی.....

- ۳۳۲.....مدل سازی منابع طبیعی برای حل مشکلات زیست محیطی
- ۳۳۴.....توصیف نقش علم داده در مدل سازی آمار فضایی
- ۳۳۵.....استفاده از آمار فضایی برای رسیدگی به مسائل زیست محیطی
- ۳۳۷.....فصل نوزدهم
- ۳۳۷.....استفاده از علم داده برای رشد تجارت الکترونیکی
- ۳۳۰.....ایجاد درکی شفاف از داده ها به منظور رشد تجارت الکترونیک
- ۳۳۱.....بهینه سازی سیستم های تجارت الکترونیک
- ۳۳۲.....بررسی جنبه های تحلیلی
- ۳۳۳.....ارزیابی اپلیکیشن های پرطرفدار تحلیل وب
- ۳۳۴.....دسترسی به تحلیل ها به منظور یافتن دستاورد
- ۳۳۵.....استفاده از تحلیل برای فعالیت
- ۳۳۶.....بررسی تحلیل ها برای حفظ کاربران
- ۳۳۷.....بررسی و آزمایش استراتژی
- ۳۳۷.....جمع بندی انواع روش های آزمایش در فرآیند رشد
- ۳۳۹.....تست دستیابی
- ۳۳۹.....آزمایش فعالیت
- ۳۴۰.....بخش بندی و هدف گذاری برای رسیدن به موفقیت
- ۳۴۱.....تقسیم بندی برای رشد سریع و راحت تر تجارت الکترونیک
- ۳۴۲.....پیدا کردن مخاطبان
- ۳۴۲.....بهینه سازی کانال های شبکه های اجتماعی
- ۳۴۳.....بخش بندی و هدف مند سازی برای حفظ کاربران
- ۳۴۳.....بخش بندی و هدف قرار دادن درآمدزایی
- ۳۴۵.....فصل بیستم
- ۳۴۵.....استفاده از علم داده برای توصیف و پیش بینی اعمال مجرمانه

۳۴۶	تجزیه و تحلیل زمانی برای پیشگیری از جرائم
۳۴۷	پیش‌بینی و نظارت بر جرائم فضایی
۳۴۷	نقشه‌برداری جرائم به‌وسیله‌ی تکنولوژی <i>GIS</i>
۳۴۸	گامی فراتر با مکان - تحلیل تخصیص
۳۴۹	استفاده از آمار فضایی پیچیده برای درک بهتر جرائم
۳۴۹	ریاضیات فضایی پیشرفته
۳۵۰	آمار توصیفی
۳۵۲	کاوش مشکلات با استفاده از علم داده در تحلیل جرائم
۳۵۳	توجه به محدودیت‌های فنی
۳۵۵	بخش ششم
۳۵۵	آشنایی با ابزار ۱۰
۳۵۷	فصل بیست و یکم
۳۵۷	ده منبع پدیده برای داده‌های باز
۳۵۹	معرفی برنامه <i>Data.gov</i>
۳۶۰	بررسی داده‌های باز کانادا
۳۶۱	آشنایی با <i>data.gov.uk</i>
۳۶۲	بررسی داده‌های دفتر سرشماری ایالات متحده
۳۶۳	معرفی داده‌های <i>NASA</i>
۳۶۳	گردآوری داده‌های بانک جهانی
۳۶۵	آشنایی با داده <i>Knoema</i>
۳۶۶	صف‌بندی با استفاده از داده‌های <i>Quandle</i>
۳۶۸	بررسی داده‌های <i>Exversion</i>
۳۶۹	نگاشت داده‌های فضایی <i>OpenStreetMap</i>
۳۷۰	منابع