



INSIGHTS^x

MDX INSIGHTS LAB

Applied Data Analytics

25 AMAZING OPEN UAE DATASETS FOR ANALYTICS

Web: <https://mdx.ac.ae/insightsx-lab>
Like us on Facebook



Middlesex
University
Dubai

CONTENTS: OPEN UAE DATASETS

Contents	2
Introduction	3
Middlesex Insights Lab: An overview	4
25 Open Datasets: Indexing	5
1 st Dataset Description	6
25 th Dataset Description	30
Contributors and Editors	31



INTRODUCTION: 25 AMAZING DATASETS IN UAE FOR ANALYTICS AND RESEARCH

MSc Data Science students and members of the Middlesex Insights Lab were asked to contribute data-sets within the context of the United Arab Emirates (UAE) that could serve as an excellent source for analytics. While there are thousands of datasets available online, very few, serve as a useful resource for running a data science project or conducting research. This document is a publication of Middlesex Insights Lab, and it aims at reducing the efforts of researchers and data science enthusiasts who wish to work on open data-sets for deriving meaningful insights.

Each dataset has been carefully chosen to ensure that it is self-sufficient to run algorithms and derive insights. While some datasets are directly from a single source, some of them have been combined from multiple sources to present one snapshot. Reading the description of the dataset, one could not only get an idea of the various attributes present but also get some ideas on Data science questions that can be cracked with this data. The format of each dataset is given below:

Format of the Dataset in this document

Title & Category

Contributor (Who curated this section)

Download Links (direct download and original source)

Description of the dataset

Attributes and Variables (all column names if structured)

Data Format (How to read this dataset)

Data Science perspective

Data Science Questions that can be cracked

Hope you enjoy this publication! All the best, with your pursuit of deriving insights.

MIDDLESEX INSIGHTS LAB: APPLIED DATA ANALYTICS

The Insights^x Lab is the Data Science hub to connect the analytics community and academic fraternity with organizations willing to derive insights from their data. The lab enables academic researchers and the data science community in the UAE to help companies understand how to better monetize their data through the development and application of new predictive models and analytical approaches.

The Insights^x Lab acts as a data translator to convert the business requirements from various firms in the UAE to data science questions that could be cracked by the crowdsourcing model (analytics community). It also provides expert analytics consulting, corporate training in data science, and knowledge sharing sessions. The following are the key offerings of the Insights^x Lab: Datathons (Data Hackathons), Use Case Nights (Experts in action), Training/Consulting (Lab experts providing services), and Data Science Meetups.

Where can you contribute?

- **Membership:** If you wish to become a member of this lab and community, please check the [“Become a Member”](#) tab for registration- Its free, and it always will be.
- **Corporate Data Partner:** If you are an organization willing to bring in a data-set for our community to crack and derive insights, please contact Dr. Krishnadas Nanath (mdxinsightslab@mdx.ac.ae) for becoming a corporate data partner.
- **Data Science Expert:** If you wish to contribute as a Data Science expert for conducting guest sessions or workshops as part of the lab, please send an email to mdxinsightslab@mdx.ac.ae.

For any other way of collaboration and contribution, please contact

Dr. Krishnadas Nanath

Head and Founder- Middlesex Insights Lab

k.nanath@mdx.ac.ae & mdxinsightslab@mdx.ac.ae

Mob: +971 554360019

The Lab promotes the [Centre for Innovation in Human Experience](#) along with Robotech Lab and Immersive VR Lab.

INDEX OF DATASETS

Category & Dataset ID	Contributor	Page No.
D1: Education: KHDA dataset	Krishnadas Nanath/Sreejith Balasubramanian	6
D2: Traffic : Accident dataset	Krishnadas Nanath	7
D3: Property: DLD Transaction dataset	Juwairiya Hussain	8
D4: Education: Sector Distribution	Ansu Vellara Anish	9
D5: Health: Diseases dataset	Eslam Hasabou	10
D6: Automobiles: Car Registration dataset	Ansu Vellara Anish/ Sreejith Balasubramanian	11
D7: Trade: Foreign Trade dataset	Josephine Mung'ala	12
D8: Traffic: Accident location dataset	Abdul Manan	13
D9: Economy: World development indicator dataset (UAE)	Moses Obeng	14
D10: Finance: Financial statistics UAE dataset	Momin Ijaz	15
D11: Economy: Dubai Inspection dataset	AKARSH ASOKAN	16
D12: Property: DLD Projects data	Abdul Manan	17
D13: Health:Health and Prevention dataset	Ibrahim Adefila	18
D14:Economy: Statistics Center UAE Dataset	Gatlin George	19
D15: Traffic: Traffic Fines	AKARSH ASOKAN	20
D16: Property: Buildings dataset	Remona Johnson	21
D17: Property: DLD Property Valuation	Remona Johnson	22
D18: Commodity Trading : Customs dataset	AKARSH ASOKAN	23
D19: Dubai Courts: Verdict dataset	Abdul Mueed Mohammed	24
D20: Import & Export: Import certificates dataset	Abdul Mueed Mohammed	25
D21: Economy: GDP Distribution dataset	Firas Hamid	26
D22: Tourism: Tour Pedia and Dubai Tourism dataset	AKARSH ASOKAN	27
D23: Transportation: RTA Metro Users dataset	Krishnadas Nanath	28
D24: Electricity and Water: DEWA dataset	Krishnadas Nanath	29
D25: Education: Dubai Culture dataset	Krishnadas Nanath	30

Dataset -1, Category: Education

TITLE: KHDA OPEN DATA-SET (HIGHER EDUCATION DATA, SCHOOL DATA)	
Contributor: Krishnadas Nanath/ Sreejith Balasubramanian	
Link: www.bit.ly/mdxkhdaopen (Data) Related: https://www.khda.gov.ae/en/opendata	
Description	A great dataset for all KHDA accredited universities census data. Data collection starts in 2013 and is updated every year. The second data-set is a good collection of private schools in Dubai including the Dubai School Inspection Bureau (DSIB) ratings each year since 2008.
Data-set 1: Higher education	For each year: University Name, Specialization, Enrolled UG students, Enrolled PG Students, Gender Distribution (M/F),
Data Format	Year: 2013-2018; Multiple Sheets; Structured
Data-set 2: Private School Data	For each year: School Name, DSIB Rating, Enrolments, Location, Latitude, Longitude, School Type, Website, Email, Year Established, Fee Structure (across grades)
Data Format	Year: 2008-2018; Multiple Sheets; Structured
Data Science Perspective:	<i>Q.1: Is there an impact of Rating on Fee structure trend over the years? (longitudinal) - School Data</i>
	<i>Q.2: Understanding data from global education, what specializations should Dubai introduce to enhance gender diversity?</i>
Hashtags	#Education #Schools #Universities #Fee #SchoolRating #longitudinal #KHDA

Dataset -2, Category: Traffic Accidents

TITLE: TRAFFIC ACCIDENT DATA (MINISTRY OF INTERIOR)	
Contributor: Krishnadas Nanath	
Link: www.bit.ly/mdxtrafficdata1 (Data File) Related (Source): http://data.bayanat.ae/en_GB/dataset/traffic-accidents	
Description	A data-set of traffic accidents in the UAE. Each row is an incident of the 2017 accident case and columns are the attributes related to the accident. Some fields are in Arabic, but the English one could be downloaded. Very important details like accident location, weather, seatbelt, intoxication, and others are shared for each accident.
Data-set: Road Accident Data UAE	For 2017: Accident ID, Date, Time, Location, Weather, Road Status, Driver Age, Gender, Occupation, Intoxication, SeatBelt Status, Insurance Company Name, Car Manufacturer Year
Data Format	Year: 2017 (others); Single Sheet; Structured (Arabic); 2000+ data points; More data available.
Data Science Perspective:	<i>Q.1: Which attributes (time, location, gender, car year and others) distinguish a fatal accident from a non-fatal one?</i>
	<i>Q.2: Can an accident be predicted to make sure there is enough support available?</i>
Hashtags	#Accident #Location #Seatbelt #Howoldisyourcar #traffic #Trafficpattern #RoadSafety

Dataset -3, Category: Property Transactions

TITLE: DUBAI PROPERTY SALE TRANSACTIONS (DUBAI LAND DPT)	
Contributor: Juwairiya Hussain	
Link: http://bit.ly/dldtransaction (Data File) Related (Source): https://www.dubaipulse.gov.ae/organisation/dld/service/dld-transactions	
Description	This dataset is about Dubai real estate transactions from 2005 until 2019. It is a comprehensive dataset of sale and mortgage transactions
Data-set: Dubai property sales	2005 to 2019: date, sale type , property type, area name , property subtype, area, actual worth, meter price
Data Format	Year:2005-2019; Single Sheet; Structured (English/Arabic);
Data Science Perspective:	<i>Q.1: Is increased pricing having an impact on sales of residential units across the years?</i>
	<i>Q.2: Are purchases more in the newly developed neighborhood?</i>
	<i>Q.3: Can we predict lower risk areas for safer investments?</i>
Hashtags	#RealEstateDubai #Dubai #DubaiInvestments #DubaiProperties

Dataset -4, Category: Education

TITLE: SECTOR WISE DISTRIBUTION OF STUDENT DATA	
Contributor: Ansu Vellara Anish	
Link: http://bit.ly/EduSets (Data) Related Source: http://data.bayanat.ae/en_GB/organization/ministry-of-education	
Description	A data set speaks about the student gender, count, grades and stages in different sector such as government and public sector of education.
Data-set: Dubai property sales	For each year: Sector Name, Zone name, Level of student, Stage of student ,Gender distribution and count of students.
Data Format	Year:2013-2019; Multiple Sheets; Structured (English/Arabic);
Data Science Perspective:	<i>Q.1: Why does the percentage of students vary largely depending on the zone?</i>
	<i>Q.2: What could be the possible reason where people choose private school over public school?</i>
	<i>Q.3.What are effective measure to be taken to enhance gender diversity in schools?</i>
Hashtags	#Education #Schools #khda #studentdata

Dataset -5, Category: Health

TITLE: HEALTH INFRASTRUCTURE AND NOTIFIABLE DISEASES IN UAE	
Contributor: Eslam Hasabou	
Link: https://bit.ly/35ESnpC (Data) Related: https://bayanat.ae/en/Categories/Health	
https://www.mohap.gov.ae	
Description	A dataset for Health infrastructure records . The number of hospitals , Pharmacies, bed counts and workforce in each Emirate for the years from 2012 to 2015 . Also includes the total patients attendance and notifiable diseases during this period
Data-set 1: Health Facilities and diseases	Health Facilities : Hospitals, Pharmacies, and Beds Workforce and attendance to health care facility Notifiable diseases and notifiable disease by age in 2015
Data Format	Year: 2012-2015; Multiple Sheets; Structured
Data Science Perspective:	<i>Q.1: why the number of Beds in the private sector is less although the attendance and number of hospitals are more ?</i>
	<i>Q.2: Some common diseases are not moving or increase taking into consideration the attendance numbers . What is the plan to overcome this problem for a better society ?</i>
Hashtags	#Health #Health_Care #Mohap

Dataset -6, Category: Automobiles

TITLE: VEHICLE REGISTRATION AND DETAILS	
Contributor: Ansu Vellara Anish/ Sreejith Balasubramanian	
Link: http://bit.ly/EduSets (Data) Related Source: https://www.dubaipulse.gov.ae/data/roads-and-cars/rta_car_registration-open?organisation=rta&service=roads-and-cars	
Description	A data set about the registration details of the vehicle sold.It gives the details of age group, type of vehicle ,brand, model and details of the vehicles.
Data-set: Dubai property sales	October 2019: Nationality; Gender; Person_birth_year; Vechile_type; Manufacture_year; Cylinders_num; cylinder_capacity; color; business_individual ; fuel_type; transmission_manual_automatic; number_of_passenger_seating_capcity; number_of_doors; brand_model; empty_weight; orgin_country_the_place_where_assembled; manufacturer_brand_name
Data Format	Year:2019; Single Sheets; Structured
Data Science Perspective:	<i>Q.1: What should be done to enhance the sale of vehicle among emirates?</i>
	<i>Q.2: Why could be the reason people prefer Japanese car over other?</i>
	<i>Q.3.What are the factors that depends on the sale of vehicle?(brand,cylinder capacity,cylinder number)</i>
	<i>Q41: Based on the personal details of the owner (Nationality, Gender, Age) can we predict their preferred vehicle type (color, fuel type, brand etc...)</i>
	<i>Q.5: Can we predict the preferred band based on the color, fuel type, engine capacity etc?</i>
Hashtags	#aautosales #vehiclesales #aautomotive #carsale

Dataset -7, Category: TRADE

TITLE: FOREIGN TRADE REPORT (FEDERAL CUSTOMS AUTHORITY)	
Contributor: Josephine Mung'ala	
Link: http://bit.ly/mdxFCAtrade2017 (Data)	
Related: https://www.fca.gov.ae/en/pages/opendata.aspx	
Description	An amazing dataset of the year to year UAE Direct foreign trade volume comparison for Jan to Dec 2016/2017. The dataset provides comprehensive imports/exports commodities, volume, value and region details.
Data-set: Foreign Trade report of the year 2017	For each year: Period(in months), Import, Non-Oil Export, ReExport, Weight, Value, Emirate, trade region, trade country, Duty & free zones
Data Format	Year: 2016-2017; Multiple Sheets; Structured (English and Arabic)
Data Science Perspective:	Q.1: Considering UAE population and diverse demographics, Is there a correlation or causation to the volumes of imports and exports by region?
	Q.2: Is the value of imports higher or lower based on the region of origin irrespective of weight?
	Q.3: In addition to the benefits of setting up a company in a UAE duty or free zone, what can be done to increase the volumes and value of trade?
Hashtags	#FCA #Trade #imports #Nonoilexports #ReExports #freezone #dutyfree #commodities #customs

Dataset -8, Category: Traffic

TITLE: DUBAI POLICE TRAFFIC INCIDENTS	
Contributor: Abdul Mannan	
Link: http://bit.ly/2qjZGmw (Data) Related: https://www.dubaipulse.gov.ae/data/dp-traffic/dp_traffic_incidents-open?organisation=dubai-police&service=dp-traffic	
(Different data set when compared to Dataset-2)	
Description	It is a data set related to the traffic accidents since August 2018 in Dubai. The best part about this data set is that the data is getting constantly updated by Dubai Police. Which can give us an opportunity to verify the preventive actions taken based on our findings.
Data-set:	For each accident: Unique Accident ID, accident Time, name of the person involved, accident place coordinates
Data Format	Year: Aug 2018-present; Single Sheets; Structured
Data Science Perspective:	<p><i>Q.1: Is there any correlation between the place and time of accidents.</i></p> <p><i>Q.2: Are there any specific places where accidents are occurring at a time of the day?</i></p> <p><i>Q.3: Since the data is being constantly updated, we can verify if the measures taken on account of Q.2, have been successful or not.</i></p>
Hashtags	#Traffic #Accidents #Dubai Police #Dubai #longitudinal

Dataset -9, Category: Economy

TITLE: DEVELOPMENT INDICATORS	
Contributor: Moses Obeng	
Link: http://bit.ly/33HaABb (Data) Related: https://data.worldbank.org/country/united-arab-emirates	
Description	The dataset encompasses all areas with regards to development in UAE; with indicators on population, population density, urbanization, GNI, and GDP.
Data-set 1: Development Indicators	Development indicators, compiled by Worldbank from officially-recognized international sources.
Data Format	Year: 1960-2018; Single Sheets; Structured; Metadata is provided; Close to 1400 economic indicators; Time series data.
Data Science Perspective:	Q.1: Which indicator needs to be actively improved to support the trend of development in UAE?
	Q.2 Which indicator has the most bearing on the growth in GDP in UAE
	Q.3 Are there significant correlations between prominent indicators and can there be a prediction model for an indicator like suicide rate?
Hashtags	#Development #WorldBankIndicators

Dataset -10, Category: Financial Statistics

TITLE: UAE GOVERNMENT FINANCIAL STATISTICS	
Contributor: Momin Ijaz	
Link: http://bit.ly/35LyLQX (Data) Related: https://www.mof.gov.ae/En/opendata/Pages/details.aspx?sid=2	
Description	The data discovery comprised of financial indicators intertwined into Revenue and Expense. Key indicators include Operating Balance Sheets, Asset Acquisitions, Incurring Cost for liabilities and overall net lending and borrowing.
Data-set:	UAE Government Finance Statistics (GFS) data for the Year 2014/2015/2016/2017/2018/2019 All files are encapsulated into one excel sheet.
Data Format	Year: 2014-2019; Single Sheets; Structured
Data Science Perspective:	<p>1. An interesting comparison of yearly and quarterly progress can create emerging trends for the future.</p> <p>2. Predicting that what should be an appropriate revenue based on the consumption / expenses incurred over a period of time.</p> <p>3. Identifying the most utilized area of expenses and the increase in the modes of revenue streams.</p>
Hashtags	#UAEfinancialStatistics #MDXDataScience

Dataset -11, Category: Economic Regulations

TITLE: DUBAI ECONOMIC DEVELOPMENT AUTHORITY INSPECTION REPORT	
Contributor: AKARSH ASOKAN	
Link: https://drive.google.com/open?id=1S4pq25x94xj10DgYafItaDi5P_XaGE-N (Data) Related: https://www.dubaipulse.gov.ae/data/ded_inspection_report-open/datafiles/7c786518-2f4e-4024-8470-e0592a10b2f9/ded-inspections?organisation=ded&service=ded-inspections#metadata	
Description	The data is comprised of inspection reports conducted by Dubai Economic Development Authority ranging from 1988 to 2019;listing the license type, inspection type, reasons for inspection and the actions taken. DED is responsible for core transactional services of business registration, enforcement of orderly business conduct and trademark protection in UAE.
Data-set:	DED inspection reports from the year 1988-2018 For each year by date: Inspection Report Number, Inspector Code,Inspection Type, License No,Inspection Reason, Result,Status etc.
Data Format	Year: 1988-2019; Single Sheet; Structured; Time-Series
Data Science Perspective:	<ol style="list-style-type: none"> 1. Over the Time period , the most repeated kind of offense and action taken? 2. Are there Repeated Violators? 3. Can a Change In the Economic Offence Dynamics over 30 years/Point of change if any ?
Hashtags	#DEDUAE #MDXDataScience

Dataset -12, Category: Property & Land

TITLE: DLD PROJECTS DATA	
Contributor: Abdul Mannan	
Link: https://drive.google.com/open?id=1nIGTfeRhle5qwBqbPvR5wy3Zf4B4qmrL (Data) Related: http://bit.ly/2VZCIm6	
Description	It is a rich dataset with vast information regarding all the DLD registered projects undertaken in Dubai since 2004 for Freehold properties.
Data-set:	For each registration we have Project ID, Project no., Project Name, Developer ID, Developer Name, Master Developer name, Master Developer ID, Project start Date, End date, Project ID type, Project Classification id, Escrow agent ID, Escrow agent name, Project Status, Project Completed Status(%age), completion date, cancellation date, Project Description, Property ID, Area ID, Area Name
Data Format	Year: Aug 2004-July 2019 (updated Constantly); Single Sheets; Structured
Data Science Perspective:	Q.1: What is the trend for the Projects getting cancelled?
	Q.2: Is there any correlation between cancelled projects and their respective developer, area, Escrow Agent or their particular combination?
	Q.3: Based on Q2 how can we mitigate the issue of projects getting cancelled?
	Q4: How many of the projects have overshooted their completion dates and what are the reasons?
Hashtags	#DLD #Housing #Realestate #Dubai #Projects

Dataset -13, Category: Health

TITLE: MINISTRY OF HEALTH AND PREVENTION (DISTRIBUTION OF DEATHS BY CAUSE OF DEATH, SEX, NATIONALITY, AND MEDICAL DISTRICT)	
Contributor: Ibrahim Adefila	
Link: https://bit.ly/2MxwOje (Data) Related: http://data.bayanat.ae/en_GB/dataset/distribution-of-deaths-by-cause-of-death-sex-nationality-and-medical-district	
Description	An insightful dataset of major diseases causing death in the United Arab Emirates across the medical districts in the seven emirates, western region and Al Ain. The data identified prevalent diseases and health problems causing deaths in the country between 2002 and 2008 defined by gender, nationality and locality demographics.
Data-set: Higher education	For each year: Year, Medical_District_AR, Medical_District_EN, Cause_of_Death_AR, Cause_of_Death_EN, Nationality_AR, Nationality_EN (Citizen/Non-Citizen), Sex_AR, Sex_EN (M/F), Value
Data Format	Year: 2002-2008; Multiple Sheets; Structured
Data Science Perspective:	<i>Q.1: To what extent does gender pose as a risk factor to terminal diseases in the United Arab Emirates?</i>
	<i>Q.2: Are residents of specific Emirates in the United Arab Emirates exposed to terminal diseases than inhabitants of other Emirates?</i>
Hashtags	#Health #PreventiveHealthcare #Healthcare #Diseases #Biostatistics #HealthData #OpenData

Dataset -14, Category: Economy

TITLE: DUBAI STATISTICS CENTER (ECONOMIC INDICATORS)	
Contributor: Gatlin George	
Link: https://bit.ly/31ucxi8 (Data File) Related: https://bit.ly/35Upt5l (Source)	
Description	A diverse dataset of the main economic variables that are used to measure economic performance by activities which used international standard industrial classification of all economic activities (ISIC 4) to classify all activities of establishments . The data covers the period of 2006 to 2017.
Data-set: Economic Indicators	For each year: Economic Activity, Year , Output In Aed , Intermediate Consumption In Aed, Value Added In Aed, Compensation Of Employees In Aed, Gross Fixed Capital Formation In Aed, Number Of Employees.
Data Format	Years: 2006-2017; Single Sheet; Structured
Data Science Perspective:	<i>Q.1: What industries are showing an improvement or decline in trends and what can be done to boost them?</i>
	<i>Q.2: What are the effects on the job market ?</i>
	<i>Q3: What are the geo-political factors affecting these industries?</i>
	<i>Q4: What factors can be predictive of a general decline in the economy and what can be done to arrest it?</i>
Hashtags	#Dubai Economy #Economic Development #Jobs #GDP

Dataset -15, Category: Traffic

TITLE: MINISTRY OF INTERIOR FINES REPORT	
Contributor: AKARSH ASOKAN	
Link: https://drive.google.com/open?id=1R5YMswDDcQrDi8YtDVHT8H6XXW27syTF (Data) Related: http://data.bayanat.ae/en_GB/dataset/fines/resource/5c9643ab-1146-4c0c-a82a-28f0da8d2470	
Description	The data is comprised of fines reports conducted by the Ministry of Interior ranging from 2012 to 2016; listing the year of incident , violation cause and the fines.
Data-set:	MOI fines report 2012-16 For each year : ID, violation type, fines
Data Format	Year: 2012-2016; Single Sheet; Structured
Data Science Perspective:	1. Over the Time period, there a trend of repeated offense?
	2. Range of fines for the kind of offense?
	3. How many were charged in the violators presence?
	4. Does the presence of the violator while fines applied, affect the repeated occurrence in any way?
Hashtags	#MOIUAE #MDXDataScience #Fines #Offense

Dataset-16, Category: Property & Buildings

TITLE: DUBAI MUNICIPALITY BUILDINGS INFORMATION DATASET	
Contributor: Remona Johnson	
Link: http://bit.ly/builingdataset (Data File) Related (Source): https://www.dubaipulse.gov.ae/data/dm-registration/dm_building_summary_information-open?organisation=dubai-municipality&service=dm-registration	
Description	An up-to-date detailed dataset with overall Dubai Municipality Building information. When a building permit application is submitted Dubai Building Permit System, the Dubai Municipality makes sure that the new buildings fulfill all governmental entities requirements' that participate in the permit procedures for approval. The information in the dataset is generated from Dubai BPS once a building permit is delivered.
Data-set	For each building: Building ID, Community Name, Permitted Date, Building Type, Number of Floors, Building Height (in meters), Total Area, Green Building (Yes or No), Building Usages, Plot Area, Completion Date, Building Status, Air Conditioned Area, Number of Lifts, Number of Indoor car parking, Number of Outdoor car parking, Community Number, Building Ownership Type, etc.
Data Format	Year: 1950 – Present (Updated daily); Single Sheet; Structured (Arabic & English)
Data Science Perspective:	Q.1: Does the height of a building depend on the community in which it is located?
	Q.2: What are the attributes that contribute to a permit cancellation?
	Q.3: Can a trend be identified in the building height over the years?
Hashtags	#Building #DubaiMunicipality #Community #Commercial #Residential

Dataset -17, Category: DLD Property Valuation

TITLE: DLD PROPERTY VALUATION DATASET	
Contributor: Remona Johnson	
Link: http://bit.ly/dldvaluationdataset (Data File) Related (Source): https://www.dubaipulse.gov.ae/data/dld-valuations/dld_valuation-open	
Description	A fresh dataset of property valuation collected by the Dubai Land Department since the year 2000. It contains comprehensive details about all property evaluations including all various attributes.
Data-set	For each Property valuation procedure: Procedure ID, Procedure Name, Procedure Year, Actual Worth, Status Code, Procedure Area, Property Type, Property Sub Type, Area Name, Actual Area and Property Total Value
Data Format	Year: 2000 – Present (Updated daily); Single Sheet; Structured (Arabic & English)
Data Science Perspective:	<i>Q.1: Over the years, how is the transformation in the value for each property type? What are the factors that could have impacted it?</i>
	<i>Q.2: Is it possible to predict the property value using regression?</i>
Hashtags	#DubaiLandDepartment #DLD #Property #Residential #Commercial

Dataset -18, Category: Commodity Trading

TITLE: FEDERAL CUSTOMS AUTHORITY DATASET ON EXPORT , IMPORT AND RE EXPORT OF COMMODITIES	
Contributor: AKARSH ASOKAN	
Link: https://drive.google.com/open?id=1HESxzHqJFBfHtQN3XO2Brmp8gbMfW-3Z (Data) Related: http://data.bayanat.ae/en_GB/dataset?q=2008+&groups=food-security&sort=score+desc%2C+metadata_modified+desc&page=1 (Multiple Sources)	
Description	A great dataset from the Federal Customs Authority on multiple commodity trading involving export, import and export quantities, trading amount, and countries involved over the period of 2008 to 2017. Commodities involved includes Oil, Animal Feed, Meat, Milk, Cereal, Water, Fruits & Vegetables and much more.
Data-set:	Commodity Trading Data over the period of 2008 – 2017; For each year : Commodity Type; Trade Type; Country, HS code, Weight and Trade Value
Data Format	Year: 2008-2017; Multiple Data Sources Mined to 1 Single Sheet; Structured
Data Science Perspective:	<p><i>Most Traded Commodity Over the Year by Weight/ Trade Value?</i></p> <p><i>Countries with the most substantial Trading relations with UAE?</i></p> <p><i>Most Imported Commodity Over the Year by weight/ Trade?</i></p> <p><i>Relative Comparisons of Export, Import , Re Export per Commodity per Country and its significance?</i></p> <p><i>Has the HS codes of any commodity impacted its chance of export/ import /re export?</i></p> <p><i>Chances of increasing the export tariff on any of the commodity based on HS code, Weight or Value Traded at?</i></p> <p><i>Chances of Alternate Import Options or reducing imports based on data if any ?</i></p>
Hashtags	#FCAUAE #import #export #commodity

Dataset -19, Category: Courts

TITLE: COURTS VERDICTS	
Contributor: Abdul Mueed Mohammed	
Link: Verdicts.csv (Data) Related: https://www.dubaipulse.gov.ae/organisation/dubai-courts?organisation=dubai-courts&service=courts-records	
Description	The dataset is comprised of the verdicts passed for different cases held in the courts sessions. It is huge data-set of 2.3 GB. It is updated on a regular basis and the update file can be downloaded from the source.
Data-set:	Courts verdicts - 2018-19 For each year: Case Serial Number, Case Subtype Code, Case Year, Decision Number, Insertion Date, Judgment Date, Last Update Date, Recitals.
Data Format	Year: 2018-19; Single Sheet; Structured; Big-data set
Data Science Perspective:	<i>Over the Time period , the most repeated kind of offense ? Different types of offense? How long does a particular category of offense take time to give a verdict? Different Verdicts possible for each offense?</i>
Hashtags	#MDXDataScience #Courts #Verdict #Bigdata #Casedata

Dataset -20, Category: Import Materials & Certificates

TITLE: CHAMBER CERTIFICATES OF ORIGIN	
Contributor: Abdul Mueed Mohammed	
Link: DCCI.xlsx (Data) Related: https://www.dubaipulse.gov.ae/organisation/dubai-chamber?organisation=dubai-chamber&service=chamber-dsi	
Description	The dataset is comprised of all the certificate of origins raised for different import related material brought into the UAE. It has important information on final destination country, mode of transportation, currency fee and dates.
Data-set:	Certificate of Origin For each year : sr_num, sr_cat_type_cd, sr_subtype_cd, x_channel, x_final_destination, x_invoice_date, x_language, type, x_transport_by, srv_fee_curcy_cd, created
Data Format	Single Sheet; Structured
Data Science Perspective:	<p><i>What is the major transport method for different countries?</i></p> <p><i>Which countries have the highest number of material imports?</i></p> <p><i>What channel does each country prefer to transact in with Dubai chambers?</i></p> <p><i>Which country has the highest imports passing through UAE?</i></p>
Hashtags	#MDXDataScience

Dataset-21: Category: Economic Indicator**GROSS DOMESTIC PRODUCT DISTRIBUTION PERCENTAGES BY ECONOMIC SECTOR***Contributor: Firas Hamid*Link Data File: <https://drive.google.com/open?id=1u7PHwA48igsPYwylQFitRZedc9woS8ZV>Related (Source): http://data.bayanat.ae/en_GB/dataset/gdp-distribution-percentages-by-economic-sector-2010-constant-prices

Description	Specific contribution of different economic sectors in total economic growth of UAE. The dataset illustrates percentage of distribution of gross domestic product(GDP) by economic activity with base year prices of 2010 in the UAE. It has data on these indicators from 2010 till 2017.
Data-set	For each row: Year, Name of the economic sector (English and Arabic) and Percentage contribution to GDP
Data-format	Single Sheet; Structured but not as time series; data from 2010-2018; Need conversion
Data Science Perspective	<ul style="list-style-type: none"> • <i>Can a forecasting be done for future years based on previous GDP percentage trend analysis?</i> • <i>Can there be a time series model for understanding the impact of one sector on other indicators?</i>
Hashtags	#MDXDataScience, #UAE Economy

Dataset -22, Category: Tourism

TITLE: TOUR PEDIA DATASET ON DUBAI TOURISM	
Contributor: AKARSH ASOKAN	
Link: https://drive.google.com/open?id=17Rht0xlpD94yIVx3mD1UxFgbGyzMYBiB (Data) Related: http://tour-pedia.org/about/datasets.html	
Description	An Amazing and Truly detailed attribute driven dataset from an Independent Source - Tour Pedia ; Consisting of Multiple Datasets that uses Data crawling to crawl data from Multiple social media sources such as Facebook, Foursquare, Google Places and Booking. The datasets contain details on accommodation, restaurants, places of interest and attractions in Dubai rated and reviewed by multiple users. This dataset could serve as an example on standard data recording and collecting formats.
Data-set:	Id, Name, Category, Sub category, Latitude, Longitude, Details and Reviews.
Data Format	By categories; Multiple Sheets ; Structured
Data Science Perspective:	<i>Influence of any of the categories such as POI or attractions that influence the usage of other categories or vice versa? Does any of the negative reviews of one category affect the other on a major scale? If so the numbers? Does the selection of Accommodations majorly influence the restaurant selections? Suggestions on increasing social media influence in the choices of categories? Improvements in existing setup and inter-connecting multiple categories that please the users.</i>
Hashtags	#DUBAITOURISM #mydubai #MDXDataScience

Dataset -23, Category: Transportation

TITLE: RTA METRO CHECK-INS	
Contributor: Krishnadas Nanath	
Link: www.bit.ly/rtamdx123 (Data) Related: https://www.dubaipulse.gov.ae/data/rta-rail/rta_metro_ridership-open	
Description	A very interesting data-set of commuters using RTA Metro in Dubai. It is a complete list of transactions that could happen using NOL Card, including check-ins , check-outs, and station names. Since there is a lot of transaction data, the files are divided by months which could be found in the source link
Data-set 1: RTA Metro events	For each transaction: Type of transaction (check-in or check-out), Date, Time, Metro Station (start, end), Metro Zone
Data Format	Year: March 2018 ; Structured; Each row is a transaction; Mix of quantitative and qualitative.
Data Science Perspective:	<i>Q.1: Are there zones that are overcrowded and under-utilized?</i>
	<i>Q.2: Can there be heat-map insights on incoming and outgoing flow of riders in a metro-station?</i>
	<i>Q.3 Can there be a prediction model for estimating the check-in or check-outs based on other attributes of transactions?</i>
Hashtags	#RTA #Metro #Commuters #Zones #NOLCard

Dataset - 24, Category: Electricity and Water

TITLE: DEWA ELECTRICITY AND WATER CONSUMPTION	
Contributor: Krishnadas Nanath	
Link: (Data File): http://bit.ly/36teeRm (Electricity) and http://bit.ly/34kZwtZ (Water) Related (Source): https://www.dubaipulse.gov.ae/data/dewa-consumption/dewa_electricity_consumption-open	
Description	This dataset is an exciting collection of electricity and water consumption listed as one account in a row. It not only provides the final consumption of electricity and water but also maps it to the community and business partners.
Data-set: Electricity and Water	Common for Electricity and Water: billing_portion; community; rate_category; consumption_period; calendar_month; contract_account; business_partner and consumption_unit; Rate category is also provided for water data.
Data Format	For October 2019: CSV format; Other months download needed from the source; Each row is a bill.
Data Science Perspective:	<i>Q.1: Based on the time series data, can there be a clustering of communities that exhibit similar trends?</i>
	<i>Q.2: Can we see a correlation between consumption and business partner?</i>
	<i>Q.3 Do communities exhibit a similar trend on water and electricity over the months in one year?</i>
	<i>Q.4 Are there communities who have demonstrated sustainable use of electricity and water. If yes, can this be replicated?</i>
Hashtags	#DEWA #Electricity #Water #Community #BusinessPartner

Dataset - 25, Category: Dubai Culture

TITLE: DEPARTMENT OF TOURISM & COMMERCE EVENT DATA	
Contributor: Krishnadas Nanath	
Link: (Data File): http://bit.ly/2Nafp0l Related (Source): https://www.dubaipulse.gov.ae/data/dtcm-registration/dtcm_e_permits_events-open	
Description	All events licensed by the Department of Tourism and Commerce Marketing (DTCM) need to be registered, and this dataset is now available online. DTCM generates an event permit using the e-permit system and this permit number along with other details are provided in this dataset.
Data-set: DTCM events in Dubai	For each event: permit_number organizer_name organizer_id event_name ticketed event_type event_subtype venue_name venue_id venue_address sub_venue_name event_start_date event_end_date event_start_time event_end_time sub_venue_capacity
Data Format	For March 2019: CSV format; Other months download needed from the source; Each row is an event; Structured
Data Science Perspective:	<i>Q.1: If a new event needs to be introduced in Dubai, which type would be the right choice given the trend of past events?</i>
	<i>Q.2: Can the capacity of a venue be optimized for a given type of event?</i>
	<i>Q.3 How does the market of events look like with big organizers and small ones?</i>
	<i>Q.4 Are there any possible groups of events to generate new market strategy?</i>
Hashtags	#DTCM #Events #Marketing #Permit #capacity #eventtime

CONTRIBUTORS AND EDITORS

MSc Data Science students and members of the Middlesex Insights Lab were the key contributors for this version of the open dataset collection. We thank the following curators for carefully bringing in the best of the open dataset in the United Arab Emirates.

Abdul Manan
Abdul Mueed Mohammed
Akarsh Asokan
Ansu Vellara Anish
Eslam Hasabou
Firas Hamid
Gatlin George
Ibrahim Adefila
Josephine Mung'ala
Juwairiya Hussain
Dr. Krishnadas Nanath
Momin Ijaz
Moses Obeng
Remona Johnson
Dr. Sreejith Balasubramanian

The document is compiled and edited by:

Dr. Krishnadas Nanath

Head & Founder – MDX Insights Lab
 Associate Professor
 Middlesex University Dubai
k.nanath@mdx.ac.ae

ANSU VELLARA ANISH Data Governance Expert-Insights x Lab AV460@live.mdx.ac.uk	FARHEEN M ANWAR Insight Community Organizer-Insights x Lab FF271@live.mdx.ac.uk
FATUNBI SEKONI Insight Outreach Coordinator-Insights x Lab sf837@live.mdx.ac.uk	MOHAMMED ABDUL MUEED Technical Expert-Insights x Lab am3411@live.mdx.ac.uk