لائحة البرامج الأكاديمية (الخاصة)

دبلوم في علوم البيانات
(Diploma in Data Science)

ماجستير في علوم البيانات
(M.Sc. in Data Science)

بكلية الدراسات العليا للبحوث الإحصائية
جامعة القاهرة
بنظام الساعات المعتمدة

2019

انطلاقا من المادة (13) بلائحة الكليه ونصها:

" يجوز للكليه ان تنظم وتستحدث برامج اكاديمية وبرامج خاصة ويحدد مجلس الكليه شروط القبول والقيد ومتطلبات الدراسة ونظم الامتحانات فى هذه البرامج والشهادات التى تمنحها الجامعة بناء على طلب مجلس الكليه"

تتقدم كلية الدراسات العليا للبحوث الاحصائية لإدارة الجامعة باقتراح استحداث برنامج الدراسات العليا الأكاديمي (الخاص) والذى يمنح الدرجتين التاليتين:

1- دبلوم في علوم البيانات (Diploma in Data Science).
2- ماجستير في علوم البيانات (M.Sc.in Data Science).

الرؤية:

أن يكون برنامج الدبلوم والماجستير فى علوم البيانات بكلية الدراسات العليا للبحوث الإحصائية جامعة القاهرة هو المصدر الأفضل لعلوم البيانات يوفر الأساس العلمي والمهارات العملية التي يطلبها سوق العمل المحلي والاقليمي.

الرسالة:

يلتزم برنامج الدبلوم والماجستير فى علوم البيانات بكلية الدراسات العليا للبحوث الإحصائية جامعة القاهرة بتقديم تعليم وبحث متميز لتخريج كفاءات ذات معرفة علمية ومهارات عملية في مجال علوم البيانات تؤهلهم لتنفيذ الحلول والاستشارات لجهات الأعمال المختلفة في السوق المحلي والإقليمى.

أهداف البرنامج:

يهدف برنامج الدبلوم والماجستير الي تأهيل وتزويد الخريجين بتعليم متميز متطور يلبي احتياجاتهم لتطوير خبراتهم وقدراتهم لاكتساب المهارات العملية والعلمية في المجالات الحديثة والمتقدمة في مجال علوم البيانات وتحليلها وذلك لتلبية احتياج سوق العمل المحلي والاقليمي ، كما يهدف هذا البرنامج إلى إعداد الطلاب وتجهيزهم على اجراء البحوث العلمية في هذا المجال.

وقد تم تصميم البرنامج وفقا للضوابط التالية والتي تتسق مع اللائحة الأكاديمية الخاصة بالكليه:

شروط عامه

1- تقدم طلبات الالتحاق للدراسة بالبرنامج فى المواعيد التى يقررها مجلس الكليه مع استيفاء الاوراق المطلوبة.
2- لمجلس الكليه بناء على توصية مجلس إدارة البرنامج – تحديد الاعداد المقترح قبولها ووضع القواعد المنظمة للاختيار من بين المتقدمين.
3- أن يسدد الطالب الرسوم الدراسية السنوية المقررة وفقا لما يحدده مجلس إدارة البرنامج ويعتمده مجلس الكليه.
4- يسدد الطلاب المصريين الرسوم الدراسية بالجنيه المصري بينما يسدد غير المصريين الرسوم بالعملة الأجنبية (الدولار الأمريكي).

قواعد عامة

ماده(1):الدرجات العلمية

تمنح جامعة القاهرة بناءً على طلب مجلس الكليه درجة دبلوم في علوم البيانات ودرجة الماجستير في تخصص علوم البيانات.

ماده(2):مجلس إدارة البرنامجين

يتولى إدارة برنامج دبلوم في علوم البيانات وبرنامج ماجستيرعلوم البيانات مجلس ادارة يتكون من العميد ووكيل الكليه للدراسات العليا والبحوث ومنسق للبرنامج وخمسه من أعضاءهيئة التدريس بالكليه من المتخصصين في الإحصاء وعلوم الحاسب.

ماده(3): نظام الدراسة بالبرنامجين ومتطلبات الحصول علي الدرجات العلمية

الدراسة باللغه الانجليزيه بنظام الساعات المعتمدة (الوحدات الدراسية)، وتعرف الساعة المعتمدة (الوحدة) بأنها وحدة قياس دراسية لتحديد وزن كل مقرر بالنسبة للمقررات الأخري، وهى تعادل محاضرة نظرية مدتها ساعة واحدة في الاسبوع ، أو درس تطبيقي، أو تدريب عملي مدتها ساعتان في الاسبوع، أو تدريب تطبيقي، أو تدريب عملي ميداني مدتها أربع ساعات في الاسبوع ، طوال الفصل الدراسي، ويشترط للحصول على الدبلوم أو الماجستير اجتياز الطالب بنجاح عدداً من الساعات المعتمدة وبياناتها كالتالي:

1- دبلوم في علوم البيانات

يشترط للحصول على دبلوم في علوم البيانات اجتياز الطالب (60) ساعة معتمدة بيانها كالتالي:

الساعات المعتمدة	البيان
15	مقررات اجبارية كمتطلبات للكليه
33	مقررات اجبارية
12	مقررات اختياريه
60	المجموع

- المدة القصوى للحصول على الدبلوم ستة فصـــول دراسية ، ويجوز بعد استطلاع رأى مجلس إدارة البرنامج وموافقة مجلس الكليه ومجلس الدراسات العليا بالجامعة زيادة المدة القصوى الى ثمانية فصول دراسية بصفة استثنائية.
- يراجع مجلس ادارة البرنامج المحتوى العلمى لكل مقرر مره كل عامين ويمكنه إجراء التعديلات المناسبة واعتمادها من مجلس الكليه وفقاً للمستجدات العالميه والتطبيقيه.
- يجوز لمجلس الكليه بناء على اقتراح مجلس ادارة البرنامج – طرح فصل دراسي صيفى مكثف لاتقل مدته عن ستة أسابيع دراسيه وبحد أقصى 9 ساعات معتمده.
- تخصص الأكواد منDS500 إلىDS599 لمقررات الدبلوم.

2- درجة الماجستير في علوم البيانات

تنقسم درجة الماجستير إلى مرحلتين:

المرحلة الاولى (تمهيدي الماجستير):

يجتاز الطالب بنجاح (**30**) ساعة معتمدة بيانها كالتالى:

الساعات المعتمدة	البيان
24	مقررات اجبارية
6	مقررات اختيارية
30	المجموع

وتخصص الأكواد من DS600 إلىDS699 لمقررات الماجستير.
المدة القصوى لدراسة المقررات في مرحلة تمهيدي الماجستير أربعة فصول دراسية ، ويجوز لمجلس الكليه بعد استطلاع رأى مجلس إدارة البرنامج ومجلس الدراسات العليا بالجامعة زيادة مدة الدراسة الى ستة فصول دراسية بصفة استثنائية.

المرحلة الثانية (إعداد رسالة الماجستير):

- بعد اجتياز الطالب لعدد 30 ساعه معتمده من المقررات الدراسية بنجاح يتقدم الطالب لتسجيل موضوع رسالة الماجستير، في موضوع يقره مجلس الكليه بعد اخذ رأى المشرف الرئيسي ومجلس إدارة البرنامج ويعتمده الأستاذ الدكتور نائب رئيس الجامعة لشئون الدراسات العليا والبحوث.
- يقوم الطالب بإعداد رسالة الماجستير في مجال علوم البيانات، ولا يجوز مناقشة رسالة الماجستير قبل مضى عام من تاريخ اعتماد تسجيلها (وتقدر بـ 18 ساعة معتمدة) ، ويجوز لمجلس الكليه بناء على توصية المشرف الرئيسي ومجلس إدارة البرنامج مد فترة التسجيل الى فصلين آخرين، وفى الحالات الاستثنائية يجوز لمجلس الكليه بناء علي توصية من مجلس إدارة البرنامج مد التسجيل لفصلين دراسيين اضافيين.
- يراعى عند تسجيل الرسالة أن يكون موضوع الرسالة ضمن الخطة البحثية التي يحددها مجلس إدارة البرنامج والمتوافقه مع الخطة البحثيه للكليه وأن يكون الطالب مستوفيا للشروط المطلوبة حسب لوائح الجامعة بالإضافة الى ما قد يقرره مجلس الكليه والجامعة.

القبول والتسجيل

ماده(4):فيما يلي شروط القبول والقيد ومتطلبات الدراسة ونظم الامتحانات والشهادات التي تمنحها الجامعة بناء علي طلب مجلس الكليه.

(أ): شروط القبول بدبلوم في علوم البيانات

1. نفس شروط القبول بدبلومات الدراسات العليا الأكاديمية الواردة بلائحة الكليه والتي تنص على:
" أن يكون الطالب حاصلاً على مؤهل عال معترف به من إحدى الجامعات والمعاهد المصرية الحكومية أو التي تم معادلتها من المجلس الأعلى للجامعات".
2. يجوز لمجلس الكليه قبول طلاب حاصلين على أحد دبلومات الكليه الأكاديمية للالتحاق بدبلوم علوم البيانات وفى هذه الحالة يعفى الطالب من دراسة المقررات التى سبق له دراستها فى المستوى 500 بشرط الا يكون قد مرعلى نجاحه فيها اكثر من عامين والايقل تقديره فى كل منها عن جيد على الأقل- وان يجتاز الطالب 48 ساعه معتمده بمرحلة دبلوم علوم البيانات ويحسب معدله التراكمي على 60 ساعه معتمده - كما يجوز قبول الطلاب الحاصلين على دبلومات أكاديمية مكافئه لدبلوم الكليه الأكاديمي ومعتمدة من المجلس الأعلى للجامعات.

ب- شروط القبول بدرجة الماجستير في علوم البيانات

يلتحق بالسنة التمهيدية لماجستير علوم البيانات الطلاب الحاصلين على أحد المؤهلات التالية:

1- دبلوم علوم البيانات بتقدير عام جيد على الأقل.
2- أحد دبلومات الكليه الاكاديمية فى الإحصاء أوعلوم الحاسب بتقدير جيد على الاقل مع تكليفه بحد أقصى (4) مقررات تكميليه يقرها مجلس الكليه بعد استطلاع رأى مجلس ادارة البرنامج.
3- أن يكون المتقدم حاصلا على أي من الدرجات الجامعية التالية (بتقدير جيد على الأقل):
بكالوريوس الهندسة - بكالوريوس الحاسبات والمعلومات- بكالوريوس التجارة شعب (الإحصاء- إدارة الاعمال - الشعبة الاكتوارية- رياضه وتأمين) - بكالوريوس العلوم شعبة (رياضيات- احصاء- رياضيات وحاسب) - بكالوريوس اقتصاد وعلوم سياسية شعبة (إحصاء) أواي درجه جامعيه مكافئه يقرها مجلس ادارة البرنامج ، ويجوز لمجلس إدارة البرنامج تحميل الطلاب من ذوى التخصصات الغير مرتبطة بعلوم البيانات بعدد (4) مقررات كحد أقصى يقرها مجلس الكليه ويتحتم اجتياز اختباراتها قبل قيد الطالب.
4- لمجلس الكليه قبول طلاب وافدين حاصلين على درجة جامعية أولى من الجامعات الاجنبية بعد معادلتها من المجلس الأعلى للجامعات لأحدى الدرجات العلمية المذكورة في الفقرة(3).ويجوز لمجلس إدارة البرنامج تحميل الطالب مواد تكميلية واجتيازها بنجاح قبل تسجيله لرسالة الماجستير.

ماده(5) :التسجيل والحذف والإضافة والانسحاب للمقررات

1- يحدد عميد الكليه مواعيد التسجيل والانسحاب والإضافة فى بداية كل فصل دراسى بناء على توصية مجلس إدارة البرنامج.
2- يتم التسجيل للطلاب فى بداية كل فصل دراسى بحد أقصى 15 ساعة معتمده وبحد أدنى 9 ساعات معتمده ويجوز لمجلس ادارة البرنامج وموافقة مجلس الكليه ان يسمح بالدراسه بالفصل الصيفى وفقا للقواعد المنظمه التى يقرها مجلس الكليه وبحد اقصى 9 ساعات معتمده.
3- لا يجوز للطالب التسجيل فى مقرر إلا بعد دراسته للمتطلبات السابقة لهذا المقرر.
4- يجوز إنسحاب الطالب من بعض المقررات وذلك قبل موعد الامتحان النهائى بثلاثة أسابيع بناء على طلبه وفقاً للقواعد المنظمة، ولا يسترد أية مصروفات بعد الاسبوع السادس من بدء الدراسة.

ماده(6):إلغاء أو ايقاف القيد

يلغى قيد الطالب فى الحالات التالية:

- الإخلال بالقواعد والنظم الجامعية المتعارف عليها.
- عدم سداد الرسوم فى المواعيد المقررة ولن يقبل عذر التخلف عن الميعاد المحدد.
- تجاوز الطالب المدة القصوى للدراسة بالدرجة (ثلاث سنوات دراسيه) دون استكمال متطلب الحصول عليها.
- عدم صلاحية رسالة الماجستير لرفض لجنة الحكم أو لثبوت سرقة علمية.
- إنقطاع الطالب عن البحث فى مرحلة إعداد رسالة الماجستير أو عدم جديته وذلك بناء على التقرير السنوى الوارد من المشرف الرئيسي.
- القيد في أكثر من دبلوم أو درجة جامعية عليا فى نفس الوقت داخل أو خارج الجامعة.

يوقف قيد الطالب فى الحاله التالية:

اذا تقدم الطالب بطلب لإيقاف القيد بحد اقصى ثلاث سنوات طوال فترة قيده وتسجيله للدرجة بالدراسات العليا وفق الشروط والقواعد المنظمة والمعمول بها فى الدراسات العليا وموافقة مجلس ادارة البرنامج ومجلس الكليه على طلبه ولايجوزإيقاف القيد بأثر رجعى أياً كانت اسبابه.

ماده (7): إعادة القيد

أ- يجوز إعادة قيد الطالب الذى تم الغاء قيده بناء على طلبه وذلك بعد موافقة مجلس الكليه وتطبق عليه اللائحة الحالية، ويجوز أن يعفى من بعض المقررات الدراسية التي سبق دراستها بنجاح وبتقدير جيد (بحد اقصى 4 للدبلوم و 2 للماجستير) ، وبشرط إلا يكون مضى على إلغاء قيده اكثر من عامين، وألا يكون مضى على دراستة اكثر من خمس سنوات.

ب- فى حالة إلغاء تسجيل رسالة الماجستير يجوز للطالب إعادة قيده بعد عام دراسي كامل ويجوز ان يكون التسجيل فى موضوع جديد ينطبق عليه شروط التسجيل الجديد أو فى نفس الموضوع الذى تم إلغائه على ألا يكون التسجيل السابق قد مضى عليه مدة عامين على الأكثر.

ضوابط الامتحانات

ماده(8) : الضوابط

- يكون الامتحان النهائي تحريريا فى جميع المقررات فيما عدا مقرر المشروع وفى حاله وجود جزء عملى أو تطبيقى فى المقرر يخصص للامتحان التحريرى 60 درجة وللجانب العملى أوالتطبيقى أو الشفهى 40 درجة.
- مدة الامتحان ثلاث ساعات ويعقد الامتحان التحريري فى نهاية كل فصل دراسي.
- النهاية العظمي للمقرر النظرى 100 درجة .
- يقيم نجاح أو رسوب الطالب فى المقررات (دبلوم – تمهيدي ماجستير) وفقاً للتقديرات التالية:

تحسب التقديرات التي يحصل عليها الطالب في كل مقرر كما يلي:

الرمز	عدد النقاط	النسبة	التقدير
أ (A)	4.00	≥ 90	ممتاز
أ- (A-)	.673	85 : < 90	
ب + (B+)	.333	80 : < 85	جيد جداً
ب (B)	.003	60 : < 80	
ج + (C +)	2.67	70 : < 60	جيد
ج (C)	2.33	65 : < 70	
د (D)	2.00	60 : < 65	مقبول
ر (F)	0.00	60 <	راسب
غ (غائب) (Abs)	0.00	-	راسب
غ م (غير مكتمل (I)	0.00	-	غير مكتمل
من (منسحب) (W)	0.00	-	منسحب
مح (محروم) Barred (Ba)	0.00	-	محروم

- يكون التقدير العام للمعدل التراكمي عند تخرج الطالب يحسب كالآتى:
  إجمالي عدد نقاط الطالب في المقررات مقسوماً على عدد المقررات
- فى حالة رسوب الطالب فى مقرر أو اكثر يسمح له بإعادة الامتحان بحد أقصى مرتين ، ويحتسب له تقدير D.

ماده(9): الحضور

يحرم الطالب من حضور الامتحان فى أى مقرر إذا كانت نسبة حضوره تقل عن 70% من مجموع الساعات لهذا المقرر ما لم يكن قد تغيب بعذر يقبله استاذ المقرر ويعتمده مجلس إدارة البرنامج مع مراعاة ما تنص عليه القواعد المقررة فى الجامعة.

ماده(10): المشروع

تنظم الدراسة بمقرر المشروع كالاتي:

1- لا يجوز للطالب التسجيل في مقرر المشروع الا بعد إجتيازه بنجاح 42 ساعه معتمده على الاقل في مرحلة الدبلوم أو 12 ساعه معتمده على الاقل في مرحلة تمهيدي الماجستير.
2- يقوم الطالب بإعداد المشروع البحثي تحت أشراف عضو هيئة تدريس أو أكثر، وتشكل لجان من أعضاء هيئة التدريس يحددها مجلس إدارة البرنامج لمناقشة المشاريع فى نهاية كل فصل التخرج، وتطبق عليه نفس نظام التقييم كباقي المواد.

المقررات الدراسية

اولاً: دبلوم في علوم البيانات

List of Courses for Diploma in Data Science

General Requirements (15 Credits)

Code	Course Title	Prere-quisite	Hours / Week			Total	Scores			Total
			Lectures	Lab	Activities		Lectures	Lab	Activities	
DS500	Mathematical Foundations	-	2	-	2	3	60	-	40	100
DS501	Statistical Methods	-	2	-	2	3	60	-	40	100
DS502	Programming Foundations	-	2	2	-	3	60	40	-	100
DS503	Introduction to Biostatistics and Demography	-	3	-	-	3	100	-	-	100
DS504	Introduction to Operations Research	-	2	-	2	3	60	-	40	100

Major Requirements (33 Credits)

Code	Course Title	Prere-quisite	Hours / Week			Total	Scores			Total
			Lectures	Lab	Activities		Lectures	Lab	Activities	
DS505	Linear Algebra	DS500	2	-	2	3	60	-	40	100
DS510	Artificial Intelligence	-	2	2	-	3	60	40	-	100
DS515	Introduction to Data Mining	-	2	2	-	3	60	40	-	100
DS520	Probability and Statistical Distributions I	DS501	2	-	2	3	60	-	40	100
DS530	Applied Statistical Inference	DS501	3	-	-	3	100	-	-	100
DS535	Decision Support Systems and Business intelligence	DS504	3	-	-	3	100	-	-	100
DS540	Practical Machine Learning	-	2	2	-	3	60	40	-	100
DS545	Introduction to Big Data Technologies	DS502	2	2	-	3	60	40	-	100
DS550	Fundamentals of Data Science	DS502	2	2	-	3	60	40	-	100
DS598	Project Preparation Skills	-	3	-	-	3	100	-	-	100
DS599	Capstone Project	DS598	3	-	-	3	100	-	-	100

Core Electives (12 Credits)

Code	Course Title	Prere-quisite	Hours / Week			Total	Scores			Total
			Lectures	Lab	Activities		Lectures	Lab	Activities	
DS555	Applied Multi-variate Analysis	-	2	-	2	3	60	-	40	100
DS560	Probability and Statistical Distributions II	DS520	3	-	-	3	100	-	-	100
DS565	Probability models for Biostatistics	DS520	3	-	-	3	100	-	-	100
DS570	Applied Regression	DS530	2	-	2	3	60	-	40	100
DS575	Data Base Systems	-	2	2	-	3	60	40	-	100
DS580	Statistical Methods for Correlated Data	DS530	2	-	2	3	60	-	40	100
DS585	Bayesian Inference	DS520	3	-	-	3	100	-	-	100
DS586	Graph Theory	DS505	2	2	-	3	60	40		100
DS587	Data Visualization	DS501	3	-	-	3	100	-	-	100
DS588	Advanced Programming for Data Science	DS502	2	2	-	3	60	40	-	100
DS589	Information Theory and Coding	-	3	-	-	3	100	-	-	100
DS590	Software Engineering	DS505	2	2	-	3	60	40	-	100
DS591	Fundamentals of Computer Vision	DS510	2	2	-	3	60	40	-	100
DS592	Data Communication and Computer Networks	-	3	-	-	3	100	-	-	100

Suggested Study Plan

All Students joining the diploma of Data Science will follow the suggested study plan for four semesters.

The suggested study plans for the four semesters are as follows:

First Semester:

Code	Course Title	Prere-quisite	Hours / Week			Total	Scores			Total
			Lectures	Lab	Activities		Lectures	Lab	Activities	
DS500	Mathematical Foundations	-	2	-	2	3	60	-	40	100
DS501	Statistical Methods	-	2	-	2	3	60	-	40	100
DS502	Programming Foundations	-	2	2	-	3	60	40	-	100
DS503	Introduction to Biostatistics and Demography	-	3	-	-	3	100	-	-	100
DS504	Introduction to Operations Research	-	2	-	2	3	60	-	40	100
Total Credit	15									

Second Semester:

Code	Course Title	Prere-quisite	Hours / Week			Total	Scores			Total
			Lectures	Lab	Activities		Lectures	Lab	Activities	
DS505	Linear Algebra	DS 500	2	-	2	3	60	-	40	100
DS520	Probability and Statistical Distributions I	DS501	2	-	2	3	60	-	40	100
DS550	Fundamentals of Data Science	DS502	2	2	-	3	60	40	-	100
	Elective	-	3	-	-	3	100	-	-	100
	Elective	-	3	-	-	3	100	-	-	100
Total Credit	15									

Third Semester:

Code	Course Title	Prere-quisite	Hours / Week			Total	Scores			Total
			Lectures	Lab	Activities		Lectures	Lab	Activities	
DS510	Artificial Intelligence	-	2	2	-	3	60	40	-	100
DS515	Introduction to Data Mining	-	2	2	-	3	60	40	-	100
DS530	Applied Statistical Inference	DS501	3	-	-	3	100	-	-	100
DS598	Project Preparation Skills	-	3	-	-	3	100	-	-	100
	Elective									
Total Credit	15									

Fourth Semester:

Code	Course Title	Prere-quisite	Hours / Week			Total	Scores			Total
			Lectures	Lab	Activities		Lectures	Lab	Activities	
DS540	Practical Machine Learning	-	2	2	-	3	60	40	-	100
DS545	Introduction to Big Data Technologies	DS502	2	2	-	3	60	40	-	100
DS535	Decision Support Systems and Business intelligence	DS504	3	-	-	3	100	-	-	100
	Elective									
DS599	Capstone Project	DS598	3	-	-	3	100	-	-	100
Total Credit	15									

توصيف المقررات
دبلوم علوم البيانات

أولاً: توصيف مقررات دبلوم علوم البيانات

Course Description: Diploma In Data Science

Required Courses

DS500: Mathematical Foundations

Mathematical Topics Fundamental to Computing and Statistics Including Trees And Other Graphs- Counting in Combinatorics - Linear Algebra- Fundamental Concepts of Calculus in One and Several Variables.

DS501: Statistical Methods

Data Analysis Including Descriptive Statistics- Data Visualization- Probability Lecture Including Basic Probability Calculations-Random Variables- Distributions- Statistical Methods Including Point and Interval Estimates-Hypothesis Testing- Regression-Probabilistic Models in Data Science Applications- Naive Bayes Classifiers and Topic Models for Text- Hidden Markov Models for Sequences.

DS502: Programming Foundations

This introduction python will teach python for data science, as well as programming in general. The course covers:

python Basics - your first program - Type - Expressions and Variables - String Operations - Python Data Structures - Lists and Tuples - Sets - Dictionaries - Python Programming Fundamentals - Conditions and Branching - Loops -Functions. Objects and Classes - Working with Data in Python - Reading files with open. Writing files with open - Loading data with Pandas - Numpy Finally, you will create a project to test your skills.

DS503: Introduction to Biostatistics and Demography

This course provides an introduction to biomedical data science with an emphasis on statistical perspectives: the process of collecting, organizing, and integrating information toward extracting knowledge from data in public health, biology and medicine. The course covers four critical aspect of data science: (i) data acquisition, common data models, and database design; (ii) data transformation and integration (wrangling); (iii) interactive and dynamic exploratory data analysis; and (iv) healthcare data visual analytics. Students will learn how to apply these ideas and create reproducible data analyses using a powerful open source tool chain (R-studio, gg-plot, GitHub, etc.). There will also be a central focus on statistical principles that guide all aspects of healthcare data analytics including the relevance of sampling schemes on potential generalizability, and the impact of both data quality/completeness and potential confounding factors on the reliability of scientific conclusions.

DS504: Introduction to Operations Research

Introduction- Problem Solving And Decision Making- Qualitative Analysis Quantitative Analysis-Model Development-Data Preparation- Model Solution-Introduction To Linear Programming- The Graphical Solution of Two-Variable Linear Programming Problems-Treatment of Special Cases- General Algebraic Approach-Solving Linear Systems Equations- Introduction To Simplex Algorithm - Sensitivity Analysis of Optimization and Other Problems- Integer Programming - The Branch-And-Bound Methods for Solving Mixed Integer Programming Problems- Transportation - Assignment - Network Models.

DS505: Linear Algebra

Topics include systems of linear equations and their solutions, matrices and matrix algebra, inverse matrices; determinants and permutations; real n-dimensional vector spaces, abstract vector spaces and their axioms, linear transformations; inner products (dot products), orthogonality, cross products, and their geometric applications; subspaces, linear independence, bases for vector spaces, dimension, matrix rank; eigenvectors, eigenvalues, matrix diagonalization. Some applications of linear algebra will be discussed, such as computer graphics, Kirchoff's laws, linear regression (least squares), Fourier series, or differential equations.

DS510: Artificial Intelligence

The Basic Ideas and Techniques Underlying the Design of Intelligent Computer Systems- The Statistical and Decision Theoretic Modeling Paradigm- Applications Ranging from Diagnosis to Game-Playing To Language Processing To Vision To Medicine.

DS515: Introduction to Data Mining

Introduction to Data Mining- Data Warehousing and OLAP- Data Preprocessing- Data Mining Knowledge Representation-Attribute-Oriented Analysis -Data Mining Algorithms-Association Rules - Classification - Prediction -Evaluating What's Been Learned -Mining Real Data-Clustering Advanced Techniques- Data Mining Software – Applications.

DS520: Probability and Statistical Distributions I

Exploring Univariate data -Introduction to Probability -Discrete Distributions- Continuous Distributions -Bivariate Data -Samples and Experiments- Estimation - Tests of Significance -Inference for Regression.

DS530: Applied Statistical Inference

An Introduction to Statistics - The Science of Collecting- Organizing- Interpreting Numerical - Descriptive Statistics- Probability Lecture- Statistical Inference.

DS535: Decision Support System and Business Intelligence

This course will demonstrate in the real environment managerial applications such as the basics of the MS SQL data mining and will provide the knowledge about the possibilities of Business Intelligence (BI) use. It will examine the BI tasks management, critical success factors of BI, Planning and analysis design and modeling design, development and implementation of information technology-based systems that support managerial and professional work, including Communications-Driven and Group Decision Support Systems (GDSS), data-Driven DSS, Model-Driven DSS and knowledge-Driven DSS.

DS540: Practical Machine Learning

Mathematical Foundations of Machine Learning- Overview of Machine Learning- Basics of Parameter Estimation - Classification Algorithms- Linear and Non-Linear Algorithms - Kernel Methods - Representation Learning and Matrix Factorization- Basics of Graphical Models- Ensemble Methods- Practical Aspects in Machine Learning- Introduction to PAC Learning- Sample Selection- Bias-Learning From Graph Data- Learning From Sequential Data.

DS545: Introduction to Big Data Technologies

The course introduces Big Data problems and associated frameworks and technologies. First the course motivates the topic using real-world big data problems. Second, it sheds light on handling big data from big data collection, to monitoring storage, analysis and reporting. Common data frameworks namely Hadoop and Map-reduce.

DS550: Fundamentals of Data Science

Introduction to the fundamentals of the data science domain. Fundamentals of the R programming language and related tools for usage in Data Science. Problem Solving using R. Basic statistics, data gathering, preparation, and analysis, data visualization, case studies, and some ethical issues.

DS598: Project Preparation Skills

Use Microsoft Excel to explore data-Use Transact-SQL to query a relational database-Create data models and visualize data -Apply statistical methods to data-Extract insights from data and create impact for companies-Use R or Python to explore and transform data-Apply data science techniques to common scenarios-Follow a data science methodology-Identify your strengths, refine your goals-Work in a team under a team leader-Build meaningful connections, meet potential employers and join a community of lifelong learners-Solve problems alongside high-achieving students with diverse backgrounds. in science, data analysis, engineering, mathematics, and more-Write a well-structured report-Present your work.

DS599: Capstone Project

The purpose of the Capstone Project is for the students to apply theoretical knowledge acquired during the Data Science program to a project involving actual data in a realistic setting. During the project, students engage in the entire process of solving a real-world data science project, from collecting and processing actual data to applying suitable and appropriate analytic methods to the problem. Both the problem statements for the project assignments and the datasets originate from real-world domains similar to those that students might typically encounter within industry, government, non-governmental organizations (NGOs), or academic research. Depending on the project's complexity, students will work individually or in small teams on a problem statement, typically specified by a faculty, industry, or governmental sponsor. The sponsor will usually be responsible for supplying the relevant data set. Research groups may propose projects. A list of possible projects will be posted early in the semester so students can align themselves with problems statements corresponding to their individual interests. Pending approval by the Course Director, students are free to design their own problem statement and construct their own data set. As the project and problem statements warrant, students may be permitted to organize into teams of two to three participants. Teams larger than three will be considered for approval on a case-by-case basis. Each project team will be supervised by the Course Director.

Elective Courses

DS555: Applied Multi-variate Analysis

Multivariate Normal Distribution-Hoteling's T2 - Test-Multivariate Analysis Of Variance (MANOVA)-Principal Component Analysis (PCA)-Factor Analysis-Models For Regression Analysis With Collinear Explanatory Variables Such As Principal Component Regression (PCR) And PLS- Analysis Of Data From Experiments With Repeated Measurements- Discriminant Analysis- Cluster Analysis.

DS560: Probability and Statistical Distributions II

Probability Axioms-Conditional Probability-The Law of Total Probability- Bayes' Theorem- Independence-Discrete and Continuous Random Variables-Multiple Random Variables-Sum of Random Variables-The Sample Mean- Sampling Distributions- Statistical Inference.

DS565: Probability models for Biostatistics

Data Types: Nominal - Ordinal and Continuous Variables - Descriptive Measures-Graphical Presentation-Randomness In Biological Observations - Statistical Thinking around Randomness-Data from Experiments and Observational Studies-Probability Concepts. Probability Distributions- Normal and Binomial-Parameter Estimation-Mean and Proportion-Quantification Of Statistical Uncertainty-Standard Error- Confidence Intervals And Hypothesis Test-Elementary Hypothesis Testing - Elementary Considerations Concerning Power And Samples Size-Labs involving interactive data analysis using statistical software. Topics Covered Include Simple And Multiple Linear Regressions- Confidence Intervals And Hypothesis Tests For Regression Coefficients- Model Selection And Model Diagnostics.

DS570: Applied Regression

Review of Simple Linear Regression - Multiple Regression With Emphasis On Lecture Of Least Squares Estimation- Residual Analysis- Model Interpretation-Transformations of Variables- Interactions- Model Selection Techniques- ANOVA-Influence Diagnostics and Multicollinearity - Box-Cox Transformations- Weighted Regression- Logistic And Poisson Regression.

DS575: Database Systems

The Fundamental Concepts of Database Systems - Data Models (ER- Relational-And Others)- Query Languages (Relational Algebra- SQL- and Others)-Implementation Techniques of Database Management Systems (Index Structures - Concurrency Control- Recovery- And Query Processing)- Management of Semi structured and Complex Data- Distributed and NoSQL Databases.

DS580: Statistical Methods for Correlated Data

General Linear (Mixed) Models for Normally Distributed Data And Generalized Linear (Mixed) Models or Non-Normally Distributed Data- Such As (Binomial-Poisson- and Gamma)- Graphical Data Exploration- Correlation Structures-Parameter Estimation -Testing / Inference- Model Selection and Interpretation-Diagnostics - Model Limitations-Modeling Will Be Done Using SAS.

DS585: Bayesian Inference

This Course Introduces Hierarchical Bayesian Statistical Methods That Enable Investigators to Combine Information From Similar Experiments- Account For Complex Spatial- Temporal- And Other Correlations- And Also Incorporate Prior Information Or Expert Knowledge (When Available) Into A Statistical Analysis- The Lecture Behind Bayesian Methods and Their Practical Implementation- and Also Compares Them With Classical (Frequentist) Methods. The Course Emphasizes Data Analysis Via Modern Computer Simulation Methods Using Win bugs and R (Free Statistical Software) That Are Introduced and Used in The Course.

DS586: Graph Theory

Set-theoretic definition of a graph, Bipartite graph, directed acyclic graph, and tournament. Matchings, Hall's Theorem and Berge's Theorem, as well as the algorithms of Prim, Dijkstra, Kruskal, and Ford-Fulkerson. Trees, connectivity and Menger's Theorem. Planarity and chromatic number. Choice of topics among: graphical probability models, dynamic programming, Bayesian Belief Propagation, and Tree width.

DS587: Data Visualization

How to Find Compelling Data Sets- Clean And Analyze The Contents And Then Present That Result Through A Variety Of Methods- Including Online Interactive Charts And Maps- Introduce Students To Data Visualization Including Both The Principles And Techniques-Students Will Learn The Value Of Visualization-Specific Techniques In Information Visualization And Scientific Visualization-How To Best Leverage Visualization Methods.

DS588: Advanced Programming for Data Science

Examine A Range of Advanced Techniques for Improving The Performance Of Python Programs- Including The- Use Of Parallel Computation And GPU Acceleration- Investigate How Python Can Be Used For Big Data Analysis Using Frameworks Such As Apache Hadoop And Apache Spark. - Advanced Topics In R Programming That Are Necessary For Developing Powerful- Robust- And Reusable Data Science Tools. - Robust Error Handling- Object Oriented Programming-Profiling and Benchmarking- Debugging- And Proper Design Of Functions

DS589: Information Theory and Coding

Lecture And Coding -Review Of Probability Lecture- Information – Entropy- Information Rate- Classification Of Codes- Kraft McMillan Inequality- Source Coding Theorem- Shannon-Fano Coding- Huffman Coding- Data Compression - Extended Huffman Coding - Joint And Conditional Entropies- Mutual Information - Channel Capacity - Differential Entropy -Block Codes And Convolutional Codes- Discrete Memoryless Channels – BSC- BEC – Channel Capacity- Shannon Limit- Asymptotic Equipartition Property.

DS590: Software Engineering

Computing Essentials - Construction Technologies - Construction tools - Requirement Analysis and Specification - Requirement fundamentals - Eliciting Requirements - Requirement validation-Software Process-Process Concept-Process Implementation - Project Planning and Tracking-Software configuration management - Evolution Process and Activities-Engineering Foundations for Software-Software Modeling and Analysis-Web designing and design software-Software Verification and Validation - Foundation, metrics, and measurement of software verification and validation - Testing - Developing secure software-Professional Practice.

DS591: Fundamentals of Computer Vision

Image acquisition, image transformations, gray level operations, histogram equalization, spatial filtering, edge detection, line and circle detection, generalized Hough transform, connected components labeling. Haar features, object detection with Adaboost, applications: face detection, open CV programming.

DS592: Data Communication and Computer Networks

Data Communications-Transmission Media- Data Encoding- Transmission Modes- Error Detection And Correction- Flow Control- Multiplexing- Switching Techniques- Routing-Networking:-Network Topologies- Protocols- Layering- Standardization- Lans- Wans & Mans- Internetworking- Management- Multicast- Continuous Media-Study Of Particular Networks And Protocols: E.G. FDDI- Ethernet- ISDN- SNMP- TCP/IP- X40- ATM.

المقررات الدراسية

ثانياً: ماجستيرعلوم البيانات

List of Courses for Preliminary Master in Data Science

Core Requirements (24 credit hours)

Code	Course Title	Prere-quisite	Hours / Week			Total	Scores			Total
			Lectures	Lab	Activities		Lectures	Lab	Activities	
DS605	Advanced Topics in Statistics	-	3	-	-	3	100	-	-	100
DS610	Advanced Machine Learning	-	2	2	-	3	60	40	-	100
DS615	Data Science and Decision Support Systems	-	3	-	-	3	100	-	-	100
DS620	Introduction to Deep Neural Nets	-	2	2	-	3	60	40	-	100
DS625	R Programming For Data Scientist	-	2	2	-	3	60	40	-	100
DS630	Advanced Bayesian Inference for Data Science	-	3	-	-	3	100	-	-	100
DS698	Ethics for Data Science and Scientific Research preparation	-	2	-	2	3	60	-	40	100
DS699	Project	DS698	3	-	-	3	100	-	-	100

Core electives (6 credit hours)

Code	Course Title	Prere-quisite	Hours / Week			Total	Scores			Total
			Lectures	Lab	Activities		Lectures	Lab	Activities	
DS611	Time Series Forecasting	-	2	-	2	3	60	-	40	100
DS612	Advanced Database Systems	-	2	2	-	3	60	40	-	100
DS613	Advanced Mathematics	-	3	-	-	3	100	-	-	100
DS614	Pattern Recognition	-	3	-	-	3	100	-	-	100
DS616	Advanced Programming in Python	-	2	2	-	3	60	40	-	100
DS617	Image Processing	-	2	2	-	3	60	40	-	100
DS618	Advanced Deep Learning	-	2	2	-	3	60	40	-	100
DS619	Calculus	-	2	-	2	3	60	-	40	100
DS621	Operations Management	-	3	-	-	3	100	-	-	100
DS622	Numerical Methods for Data Science	-	2	-	2	3	60	-	40	100
DS623	Cloud Computing	-	3	-	-	3	100	-	-	100

Suggested Study Plan

All Students joining the master program of Data Science will follow the suggested study plan for two semesters.

First Semester:

Code	Course Title	Prere-quisite	Hours / Week			Total	Scores			Total
			Lectures	Lab	Activities		Lectures	Lab	Activities	
DS605	Advanced Topics in Statistics	-	3	-	-	3	100	-	-	100
DS610	Advanced Machine Learning	-	2	2	-	3	60	40	-	100
DS615	Data Science and Decision Support Systems	-	3	-	-	3	100	-	-	100
	Elective									
DS698	Ethics for Data Science and Scientific Research preparation	-	2	-	2	3	60	-	40	100
Total Credit	15									

Second Semester:

Code	Course Title	Prere-quisite	Hours / Week			Total	Scores			To tal
			Lectures	Lab	Activities		Lectures	Lab	Activities	
DS620	Introduction to Deep Neural Nets	-	2	2	-	3	60	40	-	100
DS625	R Programming For Data Scientist	-	2	2	-	3	60	40	-	100
DS630	Advanced Bayesian Inference for Data Science	-	3	-	-	3	100	-	-	100
	Elective									
DS699	Project	DS698	3	-	-	3	100	-	-	100
Total Credit	15									

توصيف المقررات
ماجستيرعلوم البيانات

ثانيا: ماجستيرعلوم البيانات

List of Courses for Preliminary Master in Data Science

Required Courses

DS605: Advanced Topics in Statistics

Fundamentals of Linear Algebra- Matrix Analysis) - Fundamentals of Probability Lecture- Statistics -Preliminary- Fundamentals of Optimization- Matrix Factorization (Completion) - Tensor Factorization(Completion) - Unconstrained Optimization Methods - Constrained Optimization Methods -Sparse Learning (Compressed Sensing) - Feature Selection - Sub Modular Optimization - Reinforcement Learning / Dynamic Decisions.

DS610:Advanced Machine Learning

Statistical Machine Learning Lecture-- Analysis and Evaluation of Statistical Models- Supervised Learning - Artificial Neural Networks- Supervised Learning - Kernel Methods- Unsupervised Learning – Clustering- Topic Modeling- Feature Engineering- Basic Reinforcement Learning- Basic Semi-Supervised Learning.

DS615: Data Science and Decision Support Systems

Explore the current and future applications of data science as a strategic decision-making tool to achieve competitive advantage in business. With emphases on obtaining decision-making value from an organization's data assets, this course will investigate the use of data science findings to develop solutions to competitive business challenges. Through case studies, you will examine how data science methods can support business decision making and discover a range of methods the data scientist can use to get people within the organization on board with data science projects.

DS620: Introduction to Deep Neural Net

Introduction to Artificial Neural network, activation functions, Multi-layer Feedforward neural network, Backpropagation, Gradient Descent, loss Functions regularization, Dropout, model selection, and optimization, Applications using Deep learning Framework tools (TensorFlow, Kears, Torch).

DS625: R programming for Data Scientists

R programming including syntax, loops, visualization using R. Applications for Data Science using R are also used using different types of data.

DS630: Advanced Bayesian Inference for Data Science

Bayesian Concepts and Methods with An Emphasis on The Data Analysis- Model Choice -How to Conduct Inference in A Bayesian Setting- Through Posterior Means- Credible Intervals and Hypothesis Testing- The Use of The Package Rstan.

DS698: Ethics for Data Science and Scientific Research preparation

Investigate the ethical issues in computer science that ultimately also pertain to data science, including privacy, plagiarism, intellectual property rights, security, confidentiality, and many other issues.
Your study of these issues will begin broadly with a look at ethical issues in computer science at large. We will then make inferences to the narrower field of data science; we will consider ethical issues and positions. The quality and integrity of decisions and inferences based on data, and how important cases and laws have shaped the legality, if not the morality, of data science-related computing. We will use case studies to investigate these issues.

DS699: Project

The purpose of the Capstone Project is for the students to apply theoretical knowledge acquired during the Data Science program to a project involving actual data in a realistic setting. During the project, students engage in the entire process of solving a real-world data science project, from collecting and processing actual data to applying suitable and appropriate analytic methods to the problem. Both the problem statements for the project assignments and the datasets originate from real-world domains similar to those that students might typically encounter within industry, government, non-governmental organizations (NGOs), or academic research. Depending on the project's complexity, students will work individually or in small teams on a problem statement, typically specified by a faculty, industry, or governmental sponsor. The sponsor will usually be responsible for supplying the relevant data set. Research groups may propose projects. A list of possible projects will be posted early in the semester so students can align themselves with problems statements corresponding to their individual interests. Pending approval by the Course Director, students are free to design their own problem statement and construct their own data set. As the project and problem statements warrant, students may be permitted to organize into teams of two to three participants. Teams larger than three will be considered for approval on a case-by-case basis. Each project team will be supervised by the Course Director.

Elective Courses

DS611:Time Series Forecasting

Introduction- Examples- Simple Descriptive Techniques- Trend- Seasonality- The Correlogram -Probability Models For Time Series- Stationary-Moving Average (MA)- Autoregressive (AR)- ARMA And ARIMA Models-Estimating The Autocorrelation Function and Fitting ARIMA Models-Forecasting- Exponential Smoothing- Forecasting From ARIMA Models-Stationary Processes in The Frequency Domain-The Spectral Density Function- The Periodogram- Spectral Analysis-State-Space Models-Dynamic Linear Models and The Kalman Filter.

DS612: Advanced Database Systems

The Module Explores Advanced Database Systems- Their Management And Their Corporate Role. At The Heart Of Information Systems Lie Database Management Systems- Transactional Database Systems- Data Warehouses- E-Commerce Databases And Databases For Storing Complex Data. This Module Looks At The Technologies- Data Models And Policies That Such Systems Require.

DS613: Advanced Mathematics

Principal Component Analysis (PCA) -Manifold Learning and Diffusion Maps-A Nonlinear Dimension Reduction Tool- Alternative To PCA- Semi-Supervised Learning - Spectral Clustering and a Guarantee for Its Performance - Concentration of Measure and Tail Bounds in Probability Both For Scalar Variables and Matrix Variables-Dimension Reduction -Compressed Sensing / Sparse Recovery- Matrix Completion- Group Testing- Approximation Algorithms In Theoretical Computer Science - The Max-Cut Problem- Clustering on Random Graphs-Synchronization.

DS614: Pattern Recognition

The Act of Taking Raw Data and Making Decisions Based On The Categories of The Pattern - Has Applied To Such Diverse Areas As Character Recognition- Data Mining- Medical Diagnosis- Image Processing- Computer Vision- Bioinformatics- Speech Recognition- Fraud Detection-and Stock Market Prediction-Various Approaches of Pattern Recognition and Decision Making Processes- Diverse Classifier Designs- Learning Algorithms- Feature Extraction and Modeling.

DS616: Advanced Programming in Python

Data Scraping form the web, advanced techniques for improving the performance of Python programs, including the use of parallel computation and GPU acceleration, using python for big data analysis using frameworks such as Apache Hadoop and Apache Spark.

DS617: Image Processing

Image Representation- Convolution Filters. 2D And 3D Signal Processing- Fourier Transforms- Image Sampling and Resembling- Gray Scale Operations- Segmentation- Morphology- Thinning-Edge Detection- Image Restoration- Image Compression.

DS618: Advanced Deep Learning

Different Deep Neural Network Architectures, Convolutional Neural networks, Recurrent Neural Network (RNN), LSTM- Deep Unsupervised Learning (Deep Belief Network), Autoencoder, Deep Reinforcement Learning- Applications to Problem Domains like Speech Recognition and Computer Vision.

DS619: Calculus

Multi-Dimensional Analysis- Concepts of Vector Calculus - Vector Field- Functions of Several Variables- Partial Derivatives - Line-Integrals- Multi-Variable Integrals.

DS621: Operations Management

Production – Management – Operations – Management – Operations-Strategy-Forecasting - Location Strategies - Facility Layout - Planning and Analysis -Total Quality Management - Business Process Modeling - Project Management - Planning Process - Project Implementation - Control and Closur - Aggregate Planning-Supply Chain Management - Operations Scheduling - Value Engineering – Just - In-Time.

DS622: Numerical Methods for Data Science

Error Analysis -Systems of Linear and Nonlinear Equations- Eigen value Problems-Least Square Methods-Fast Fourier Transform- Ordinary (And Partial) Differential Equations and their Applications.

DS623: Cloud Computing

Theoretical Models and Definitions As Well As Practical Aspects of Cloud Computing- The Fundamentals of Hardware Virtualization and Its Role in Construction of Data Centers - Autonomic Management of Data Centers-Development of Applications (Services) for This Type Of Infrastructure- In The Laboratory Part Some of The Theories and Techniques Discussed in The Theoretical Part Are Put Into Practice.