Программа учебной дисциплины
«Основы анализа и визуализации данных для медиа»

Утверждена
Академическим советом ОП
Протокол №_____ от _____._____.20_____

Разработчик	Рогович Татьяна Владимировна, Департамент больших данных и информационного поиска
Число кредитов	3
Контактная работа (час.)	38
Самостоятельная работа (час.)	76
Курс, Образовательная программа	4 (Б) курс, Реклама и связи с общественностью
Формат изучения дисциплины	Без использования онлайн курса

1. Цель, результаты освоения дисциплины и пререквизиты

Цели:

1. Развитие у студентов навыка чтения, интерпретации и оценки качества анализа и представления количественных данных
2. Освоение инструмента визуализации количественной информации
3. Формирование навыка оценки качества визуализации данных и ее корректности

Планируемые результаты обучения (ПРО):

1. Различать и интерпретировать базовые описательные статистики, виды переменных
2. Интерпретировать результаты простых экспериментальных исследований и опросов
3. Считать базовые описательные статистики (меры центральности, разброса и т.д.)
4. Создавать визуализации количественных данных с помощью Tableau
5. Выбирать корректный вид визуализации для данных
6. Знать главные когнитивные особенности восприятия визуальной информации и опираться на них при создании визуализаций
7. Различать плохие примеры визуализации данных и уметь обосновывать свое мнение

8. Замечать ошибки, допущенные при визуализации данных (искажение трендов и информации и т.д.) и избегать их при создании своих визуализаций

2. Содержание учебной дисциплины

Тема (раздел дисциплины)	**Объем в часах** лк см/пр onl/ср	**Планируемые результаты обучения (ПРО), подлежащие контролю**	**Формы контроля**
1. Описательная статистика и статистика выводов.	10 6 30	1. Различать и интерпретировать базовые описательные статистики, виды переменных 2. Интерпретировать результаты простых экспериментальных исследований и опросов 3. Считать базовые описательные статистики (меры центральности, разброса и т.д.)	КР1, Проект.
2. Визуализация данных: особенности восриятия, лучшие практики и ошибки	6 4 16	1. Выбирать корректный вид визуализации для данных 2. Знать главные когнитивные особенности восприятия визуальной информации и опираться на них при создании визуализаций 3. Различать плохие примеры визуализации данных и уметь обосновывать свое мнение 4. Замечать ошибки, допущенные при визуализации данных (искажение трендов и информации и т.д.) и избегать их при создании своих визуализаций	ДЗ1, Проект.
3. Практика по визуализации данных в Tableau	0 12 30	1. Создавать визуализации количественных данных с помощью Tableau 2. Выбирать корректный вид визуализации для данных	КР2, Проект.
Часов по видам учебных занятий:	16 22 76		
Итого часов:	114		

Содержание разделов дисциплины:

1. **Описательная статистика и статистика выводов.**
   Кейсы применения анализа данных и визуализаций в современных редакциях. Источники данных (открытые и закрытые базы, скрейпинг данных из всемирной сети, Google Trends, краудсорсинг и онлайн-эксперименты). Введение в статистику. Понятия выборки и генеральной совокупности. Методы эксперимента и наблюдения. Главные вопросы при работе с данными: как были собраны данные, насколько надежна информация. Введение в наблюдательные исследования.

Оцениваем качество выборки (случайная или неслучайная выборка, методы случайной выборки). Что нам говорят опросы? Виды переменных (категориальные, порядковые, непрерывные, дискретные). Распределения. Меры центральной тенденции и разброса. Корелляция и причино-следственная связь. Когда мы можем обобщать статистики выборки на генеральную совокупность? Тестирование гипотез и виды ошибок. Статистическая значимость. Введение в экспериментальные исследования. Интерпретация результатов экспериментальных исследований на основе оценки качества эксперимента, выборки, анализа статистических результатов. A/B тестирование. Этические аспекты экспериментальных исследований.

2. **Визуализация данных: особенности восриятия, лучшие практики и ошибки**
   Особенности восприятия визуальной информации. Основные принципы хорошей визуализации и представления данных: как отличить хорошую визуализацию от плохой, как не ввести потребителя медиа в заблуждение и не обмануться самому. Анализируем таблицы и графики - какие виды данных какой требуют презентации? Как правильно интерпретировать визуально представленную информацию. Основные виды графиков и особенности их применения (круговая диаграмма, гистограмма, тренд, линейный график, диаграмма рассеяния, пузырьковый график и т.д.). Основы сетевого анализа и интерпретация визуализаций сетей.
3. **Практика по визуализации данных в Tableau**
   Инструмент для визуализаций и аналитики Tableau: особенности интерфейса, функционал, импорт данных. Базовая работа с данными (группировка, преобразование, создание новых переменных). Разбираем на практике создание разных видов графиков и их экспорт. Создание дашбордов и визуальных историй.

3. Оценивание

- **КР1**, Не блокирующее, Контрольная работа
  Контрольная работа по разделу 1 (описательная статистика и статистика выводов). Контрольная работа проводится во время семинара.
- **ДЗ1**, Не блокирующее, Выступление с презентацией
  Презентация плохой визуализации и рекомендации по ее улучшению. Десятиминутные презентации в парах во время семинара по предварительной записи. Преподаватель оставляет за собой право назначать отвечающего из пары во время презентации.
- **КР2**, Не блокирующее, Контрольная работа
  Контрольная в tableau по визуализации данных. Работа проводится во время семинара.
- **Проект**, Не блокирующее, Экзамен (устный)
  Финальный проект по визуализации и анализу данных с устной защитой на данных по выбору.

Формула округления: Стандартное арифметическое округление

Вид формулы оценивания: Линейная
Формула оценивания:

Окончательная оценка = Округление(0.6 * (КР1 + КР2)/2 + 0.15 * ДЗ1 + 0.25 * Проект)

Способ округления результирующей оценки по учебной дисциплине: арифметический.

Академическая этика
Преподаватель оставляет за собой право устроить устную защиту любой из форм контроля.
Если во время проверки контрольной работы, проектного задания или домашнего задания выявлен факт нарушения академической этики, то студент получает оценку «0» за данную работу. Работа студента, предоставившего работу для списывания, тоже аннулируется. Преподаватель оставляет за собой право применить дисциплинарное взыскание к обоим студентам.

Если во время написания контрольных работ студент нарушает правила проведения контрольной работы, то студент получает за эту работу оценку «0». К нарушению правил относятся: списывание, коммуникация с другими студентами, использование телефона/умных часов/планшета и т.д., использование материалов, не оговоренных преподавателем.

Проектная работа сдается к дедлайну перед экзаменом (устная защита). Проектные работы, сданные после дедлайна, не принимаются. Студенту ставится неявка на экзамен при отсутствии сданной вовремя проектной работы или при неявке на устную защиту (даже при наличии сданной работы).

Если студент отсутствовал во время контрольной работы или экзамена или не смог вовремя сдать домашнее задание по причине болезни или по другим уважительным причинам, подтвержденных учебным офисом, он получает право на пересдачу контрольной работы, пересдачу экзамена или продление дедлайна. В других случаях отсутствие студента считается неявкой и он получает оценку «0» за контрольную работу или презентацию.

4. Примеры оценочных средств

Задание к итоговому проекту будет похоже на задание прошлого года. Возможно уточнение критериев оценки и формулировок.
https://docs.google.com/document/d/1ycbgqgiHLCnRpCqQtymDCWMM-JKzb4sEbf_W...

5. Ресурсы

5.1. Рекомендуемая основная литература

п/п	Наименование

1	Толстова Ю. Н. Анализ социологических данных: Методология, дескриптивная статистика, изучение связей между номинальными признаками: Учеб.пособие для вузов / Ю. Н. Толстова. – М.: Научный мир, 2000. – 352 с. - ISBN 5-89176-086-X.

5.2. Рекомендуемая дополнительная литература

п/п	Наименование
1	Tableau official free training videos

5.3. Программное обеспечение

п/п	Наименование	Условия доступа/скачивания
1	Microsoft Windows 7 Professional RUS Microsoft Windows 8.1 Professional RUS Microsoft Windows 10	*Из внутренней сети университета (договор)*
2	Microsoft Office Professional Plus 2010	*Из внутренней сети университета (договор)*
3	Tableau Public	Бесплатная версия программного продукта Tableau с ограниченными возможностями https://public.tableau.com

5.4. Профессиональные базы данных, информационные справочные системы, интернет-ресурсы (электронные образовательные ресурсы)

п/п	Наименование	Условия доступа/скачивания
	Профессиональные базы данных, информационно-справочные системы	
1	Электронно-библиотечная система Юрайт	URL: https://biblio-online.ru/
2	Сайт соревнований по машинному обучению kaggle.com	Открытый ресурс, доступ к большому количеству наборов данных
	Интернет-ресурсы (электронные образовательные ресурсы)	
1	Открытое образование	URL: https://openedu.ru/

5.5. Материально-техническое обеспечение дисциплины
Учебные аудитории для лекционных по дисциплине обеспечивают использование и демонстрацию тематических иллюстраций, соответствующих программе дисциплины в составе:
- ПЭВМ с доступом в Интернет (операционная система, офисные программы, антивирусные программы);
- мультимедийный проектор с дистанционным управлением.
Учебные аудитории для семинарских и самостоятельных занятий по дисциплине оснащены ПЭВМ, с возможностью подключения к сети Интернет и доступом к электронной информационно-образовательной среде НИУ ВШЭ.

Компьютерные классы оборудованы ПЭВМ с доступом в Интернет, операционными системами и программным обеспечением, необходимыми для освоения дисциплины. При необходимости допускается замена оборудования его виртуальными аналогами.

6. Особенности организации обучения для лиц с ограниченными возможностями здоровья и инвалидов

В случае необходимости, обучающимся из числа лиц с ограниченными возможностями здоровья (по заявлению обучающегося) а для инвалидов также в соответствии с индивидуальной программой реабилитации инвалида, могут предлагаться следующие варианты восприятия учебной информации с учетом их индивидуальных психофизических особенностей, в том числе с применением электронного обучения и дистанционных технологий:

6.1.1. *для лиц с нарушениями зрения:* в печатной форме увеличенным шрифтом; в форме электронного документа; в форме аудиофайла (перевод учебных материалов в аудиоформат); в печатной форме на языке Брайля; индивидуальные консультации с привлечением тифлосурдопереводчика; индивидуальные задания и консультации.

6.1.2. *для лиц с нарушениями слуха:* в печатной форме; в форме электронного документа; видеоматериалы с субтитрами; индивидуальные консультации с привлечением сурдопереводчика; индивидуальные задания и консультации.

6.1.3. *для лиц с нарушениями опорно-двигательного аппарата:* в печатной форме; в форме электронного документа; в форме аудиофайла; индивидуальные задания и консультации.