

§7 МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ НОВЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Москвичев А.М., Ипатов Ю.А. ■

ВИЗУАЛИЗАЦИЯ СТАТИСТИЧЕСКИХ ДАННЫХ С ЭЛЕМЕНТАМИ ГИС ТЕХНОЛОГИЙ НА ОСНОВЕ GeoFlow

Аннотация: В задаче визуализации информации возникает проблема представления в наглядной и приемлемой форме результатов исследования. Традиционные инструменты в этой области в недостаточной мере способны решить задачу визуализации. Так для визуализации данных имеющих привязку в ГИС необходимы временные, ресурсные затраты и как правило, подготовка специалистов. В тоже время, инструмент GeoFlow, позволяет быстро и эффективно визуализировать данные, а также проводить их анализ, накладывая слои на картографическую основу. Целью данного исследования является выявление потенциальных возможностей и преимуществ исследуемого программного модуля, а также возможность применения её на практике при анализе многопараметрических геопространственных данных. В работе исследуются модельные многопараметрические статистические данные, имеющие геопространственную привязку, а также современные методы их визуализации. Научная новизна исследования заключается в постановке и решении задач визуализации информации её анализа имеющих пространственную привязку и позволяет наглядно отобразить изменения, тенденции и закономерности. Визуализация данных с помощью данной программы позволяет конструировать динамические сцены, которые повышают информативность и увеличивают степень понимания и, несомненно, дают преимущество перед другими видами визуализации. Добавляя возможности 3D-визуализации данных на карте в виде различных диаграмм с возможностями анимации, которые отсутствуют в таких современных ГИС как ArcGis, MapInfo и др.

Ключевые слова: Визуализация данных, ГИС технологии, геопривязка, GeoFlow, картографическая основа, геостатистика, анализ информации, методы визуализации, многопараметрические данные, программы визуализации

Введение

Область компьютерной графики позволяет наглядно представить исследуемый массив данных и визуализировать в удобной и наглядной форме. С развитием информационных технологий требования к визуализации и преставлению информации постоянно увеличиваются. Результаты визуализации, как правило, интегрированы в информационные системы и системы поддержки принятия решений.

Как правило, в задаче визуализации данных возникает проблема представления в наглядной форме эмпирических измерений или результатов теоретического исследования. Традиционные инструменты в этой области – графики и диаграммы – в недостаточной мере способны решить задачу визуализации, когда возникает необходимость изобразить зависимости трех и более величин [1-3].

С другой стороны, существует мощнейший инструмент изображения информации, привязанной к географической карте, на сегодняшний день это быстро развивающийся арсенал ГИС-технологий. Одним из перспективных направлений исследований в этом направлении является пространственный анализ данных [4-6]. Так для освоения и использования современных ГИС систем необходимы временные и ресурсные затраты для подготовки квалифицированного пользователя. Однако, созданный инструмент GeoFlow (Microsoft Power Map Preview for Excel), позволяет быстро и эффективно визуализировать данные, а также проводить их анализ, накладывая слои на картографическую основу[7].

Целью данного исследования является выявление потенциальных возможностей и преимуществ заявленного инструментария, а также возможность применения её на практике при анализе многопараметрических геопространственных данных.

Исходя из цели работы, поставлены следующие задачи:

- выявить особенности и возможности надстройки GeoFlow (Microsoft Power Map Preview for Excel);
- провести анализ статистических данных с применением GeoFlow (Microsoft Power Map Preview for Excel);
- определить «плюсы» и «минусы» разработки и возможные перспективы её использования.

Основные особенности и возможности GeoFlow

На сегодняшний день существует довольно большое количество средств обработки и визуализации научных данных. Для офисного пакета Microsoft разработана надстрой-

Визуализация статистических данных с элементами гис технологий

ка, которая позволяет пользователям создавать динамические презентации. Так, модуль GeoFlow предоставляет возможность совместить полученные в Excel аналитические данные с картографической основой [8]. Это нововведение позволяет облегчить анализ имеющей географическую составляющую информации.

Изначально программа была предназначена для автоматизированного поиска закономерностей при анализе научных данных, привязанных к определенным точкам в пространстве (при использовании больших объемов астрономической и геологической информации). GeoFlow для Excel предлагает аналогичные возможности всем пользователям, которым необходимо увидеть или продемонстрировать взаимосвязи между данными, представленными в формате Excel с привязкой к месту или времени.

Надстройка позволяет пользователям самостоятельно настраивать несколько слоев отображения данных (до миллиона строк) на трехмерной карте, включать и выключать эти слои, настраивать вариант визуализации слоев в виде столбцов данных, окружностей, цветовых карт, а также просматривать как данные менялись с течением времени.

С помощью Geoflow можно привязывать данные Excel к местности, исследовать данные в территориальном и хронологическом разрезе, а также выполнять другие операции в интерактивном режиме. Этот модуль открывает широкие возможности анализа: с его помощью можно находить закономерности, которые трудно заметить в традиционных «плоских» цифровых таблицах и диаграммах. Таким образом, эта надстройка добавляет редактору таблиц еще больше возможностей обработки данных и бизнес-аналитики.

Для того, чтобы воспользоваться инструментом, предъявляются следующие системные требования [9]:

- Microsoft Office Professional Plus 2013 или Office 365 ProPlus;
- Пакет GeoFlow (Microsoft Power Map Preview for Excel);
- Microsoft .NET Framework 4.0;
- Windows 7/8 или Windows Server 2008R2;
- Постоянный доступ к сети Интернет (для работы с картами Bing);
- Данные с географической привязкой (широта, долгота; почтовый индекс; название страны, города, региона; адрес).

Доступная версия GeoFlow (Microsoft Power Map Preview for Excel) имеет статус предварительной. Загрузить дополнение можно на сайте Microsoft.

После установки надстройки, её можно найти на ленте «ВКЛАДКА» табличного процессора Microsoft Excel (рис. 1)

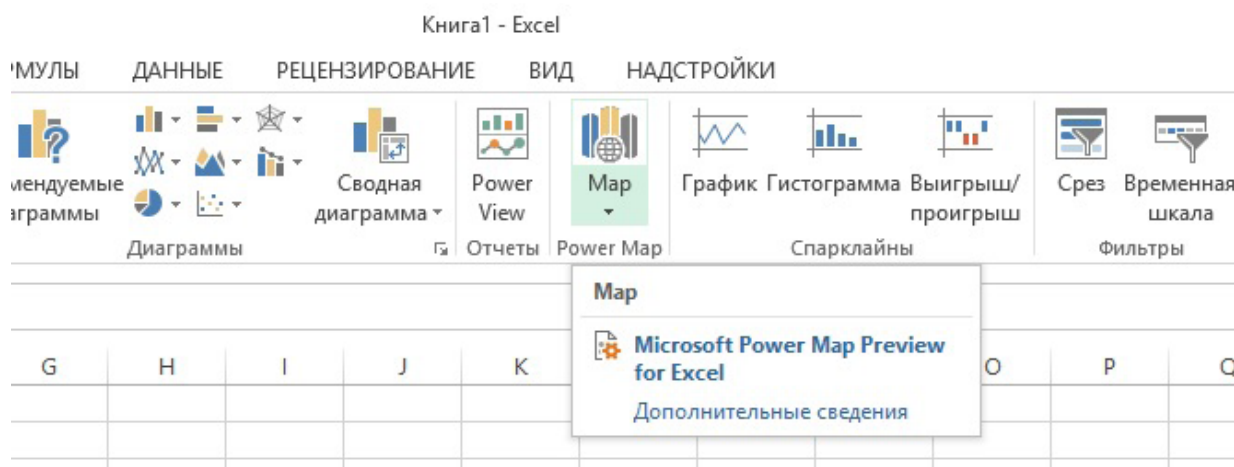


Рис.1. Microsoft Power Map Preview for Excel в табличном процессоре.

В целом благодаря GeoFlow можно:

- преобразовать данные в трехмерный вид на картах Bing. Надстройка поддерживает введение более одного миллиона строк информации из книги Excel, с Excel Data Model или PowerPivot;
- видеть новые закономерности. Сочетание географических данных и изменений во времени позволяет наблюдать тенденции, которые неочевидны при анализе обычных таблиц и гистограмм;
- делиться полученной информацией. Пользователи могут создавать видеоролики и управляемые «туры», которыми можно поделиться с широкой аудиторией.

Еще одна возможность инструмента – возможность записи последовательности переходов по карте и включения/выключения слоев данных с целью последующей демонстрации.

Специалисты Microsoft постоянно работают над расширением возможностей отображения данных Excel. Например, информацию из редактора таблиц можно наглядно показать в презентации PowerPoint. А новый Excel 2013 умеет быстро форматировать таблицы и сам подсказывает лучшие способы визуализации введенной в них информации.

Анализ статистических данных с использованием GeoFlow

Для того чтобы продемонстрировать работу надстройки, в работе использовались статистические данные с сайта Федеральной службы государственной статистики. В качестве примера возьмем статистические показатели численности экономически

Визуализация статистических данных с элементами гис технологий

активного населения, численности безработных и численности населения в целом, в тысячах человек по субъектам Российской Федерации за 2012 год [10].

Необходимо отметить, что надстройка «не понимает» таблицы с сайта Росстата, поэтому они требуют дополнительного редактирования и преобразования.

Интерфейс надстройки (рис. 2) напоминает конструктор презентаций в Microsoft Power Point, за исключением того, что каждый слайд не является статичным, а представляет собой подобие видео-ролика. Язык интерфейса – английский, также как и язык картографической основы Bing. В любой момент, нажав кнопку “Play Tour”, можно посмотреть результат визуализации в виде динамического ролика.

Клавиша “Add Scene” позволяет создать новый ролик / слайд.

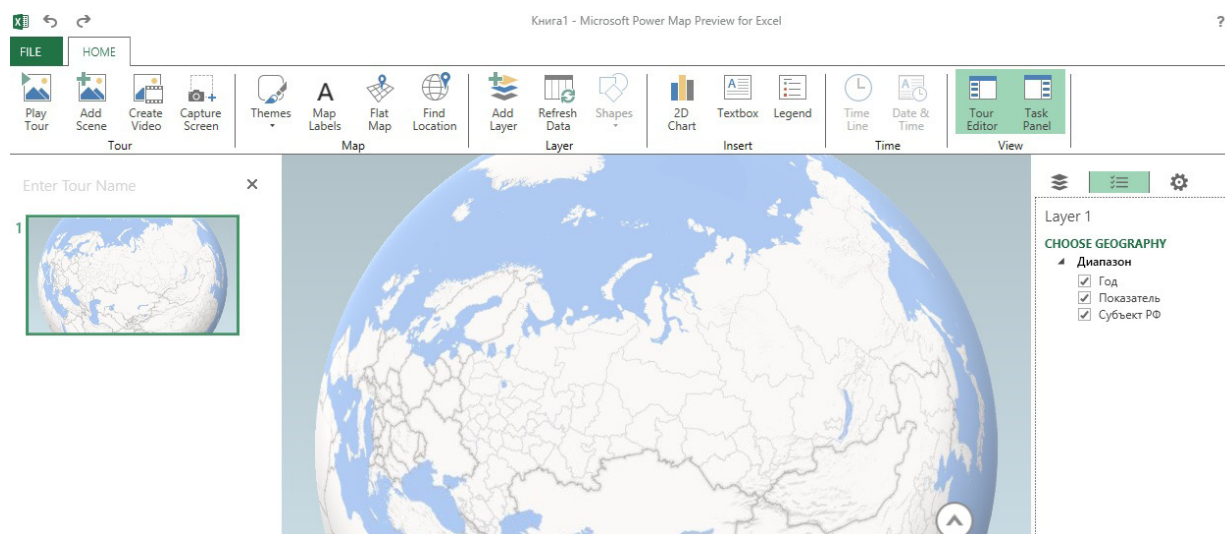


Рис. 2. Интерфейс надстройки GeoFlow (Microsoft Power Map Preview for Excel)

Данные для анализа привязываются многослойно, имеется возможность подключать / отключать слои.

С помощью мышки можно поворачивать земной шар таким образом, чтобы видимой частью стала Россия. Для изменения угла обзора, необходимо удерживать клавишу «Alt», во время перемещений мышью.

Для начала работы среди данных выбирается тот тип данных, который имеет географическую привязку. В нашем случае это «Субъект РФ». Надстройка понимает привязку по географическим координатам, наименованиям географических объектов, адресу, почтовому индексу. То есть в поле «Geography and map level» в нашем случае необходимо указать “State / Province” (рис. 3).

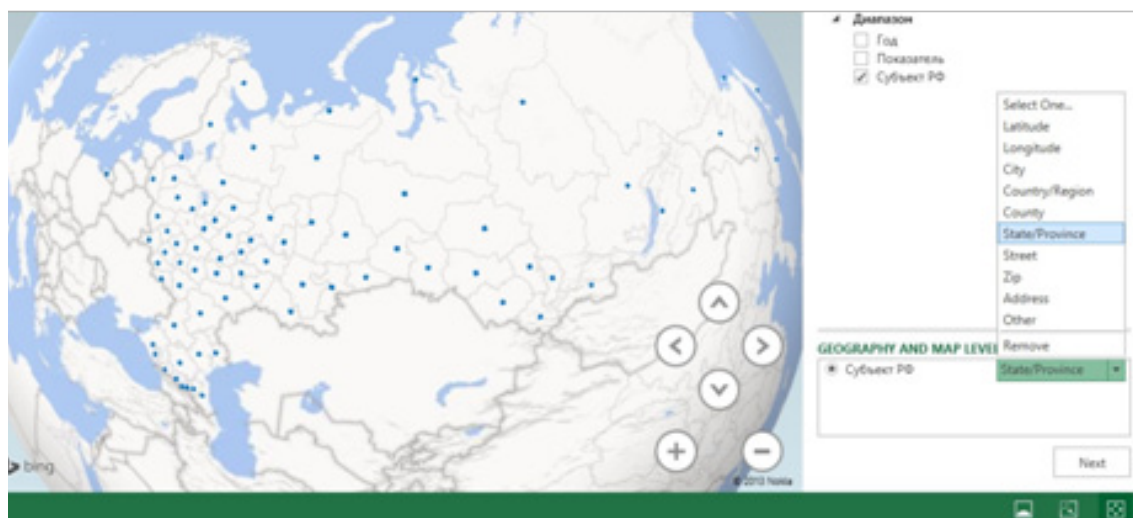


Рис.3. Выбор слоя с географической привязкой

Стоит отметить, что в документации к надстройке написано, что она воспринимает названия объектов, написанных только на английском языке. Но, как ни странно, субъекты РФ, написанные на кириллице, программа тоже нашла. В результате операции служба Bing распознала 97 % наименований субъектов РФ. Служба не распознала г. Москву и г. Санкт-Петербург (рис. 4)

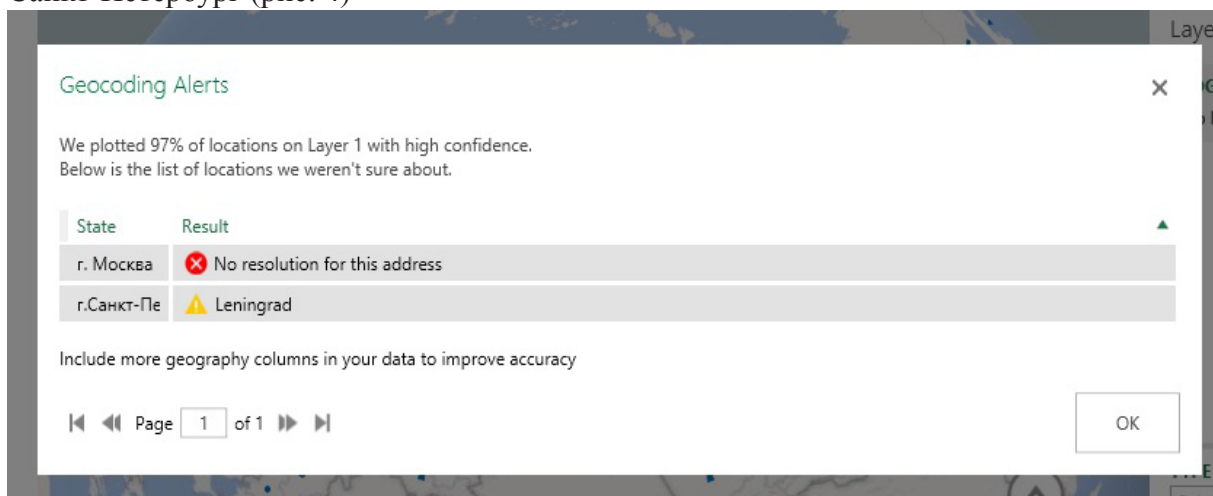


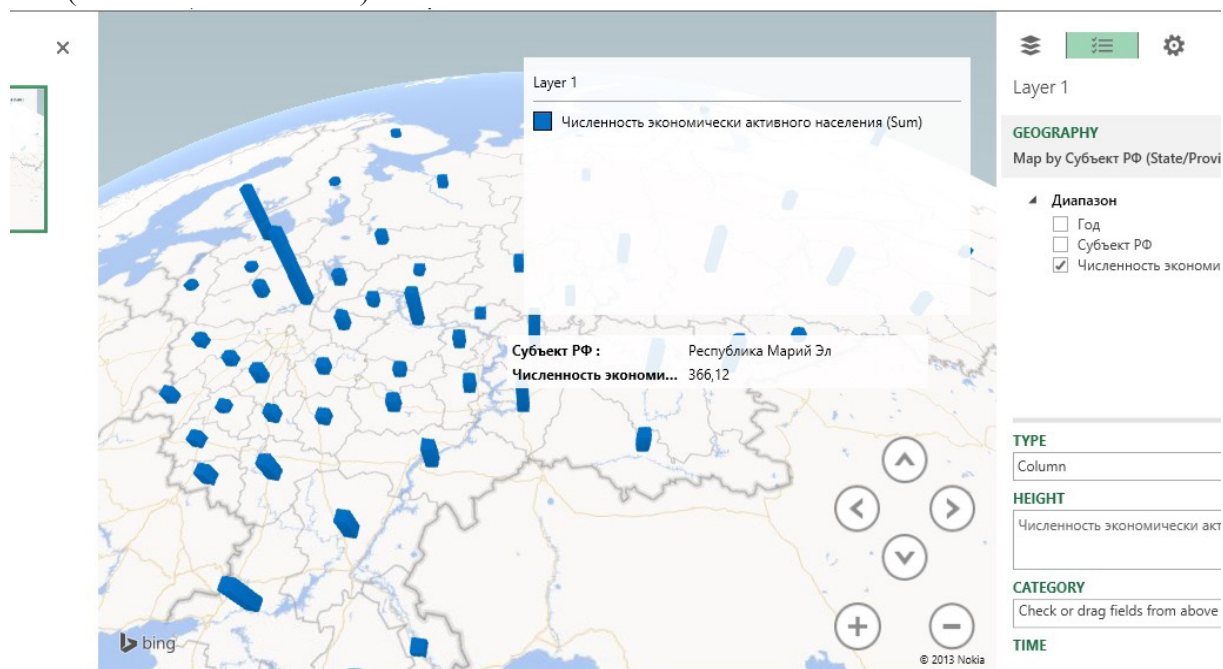
Рис. 4. Нераспознанные субъекты РФ.

Далее необходимо указать данные, необходимые для привязки. В нашем случае – это поле «Численность экономически активного населения». Можно выбрать четыре различных вида отображения данных: столбчатый, пузырьковый, термальный (тепловая карта) и региональный. Последний в зависимости от численной характеристики показав

Визуализация статистических данных с элементами гис технологий

теля окрашивает субъект либо в более бледный цвет, либо в более насыщенный. Чтобы изменить вид отображения, в правой панели в нижнем окне необходимо выбрать из выпадающего списка «Type», один из трех графиков. Для столбчатого и пузырькового вида отображения можно установить форму столбцов / пузырьков, настроить их ширину, высоту и т.д.

Имеется возможность посмотреть, как изменяются статистические данные с течением времени. Для этого необходимо нажать на кнопку «Play» на временной шкале, находящейся внизу экрана. Надстройка начнет отображение данных с самого первого дня, для которого существуют данные, и постепенно аккумулировать их с определенным шагом (шаг можно изменять).



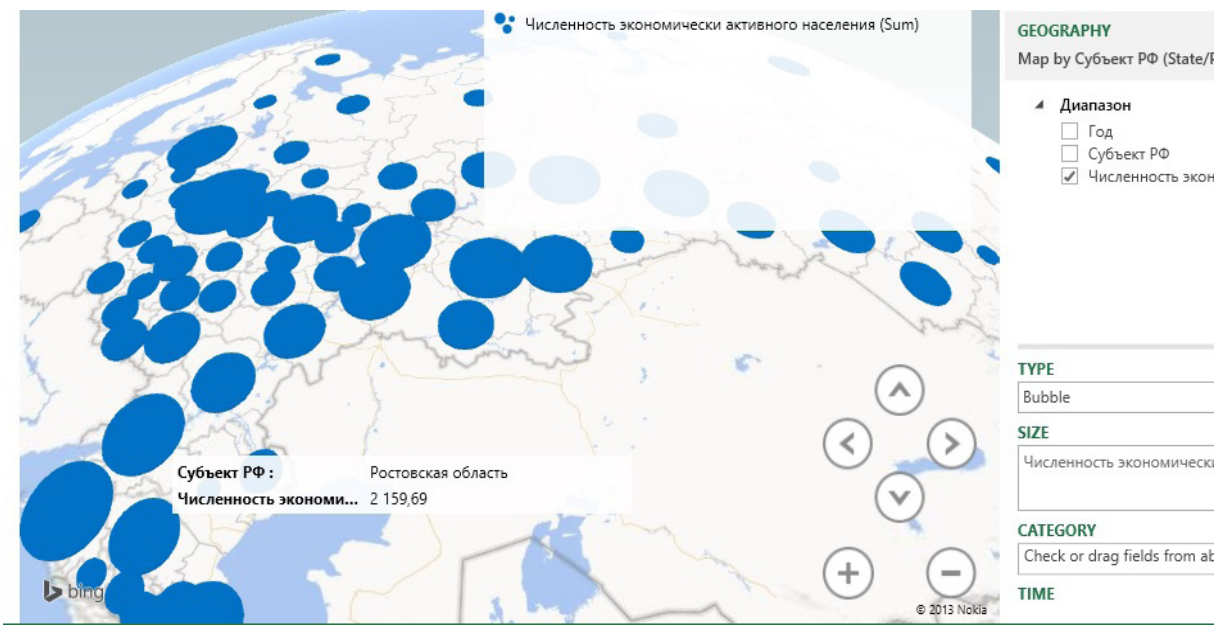
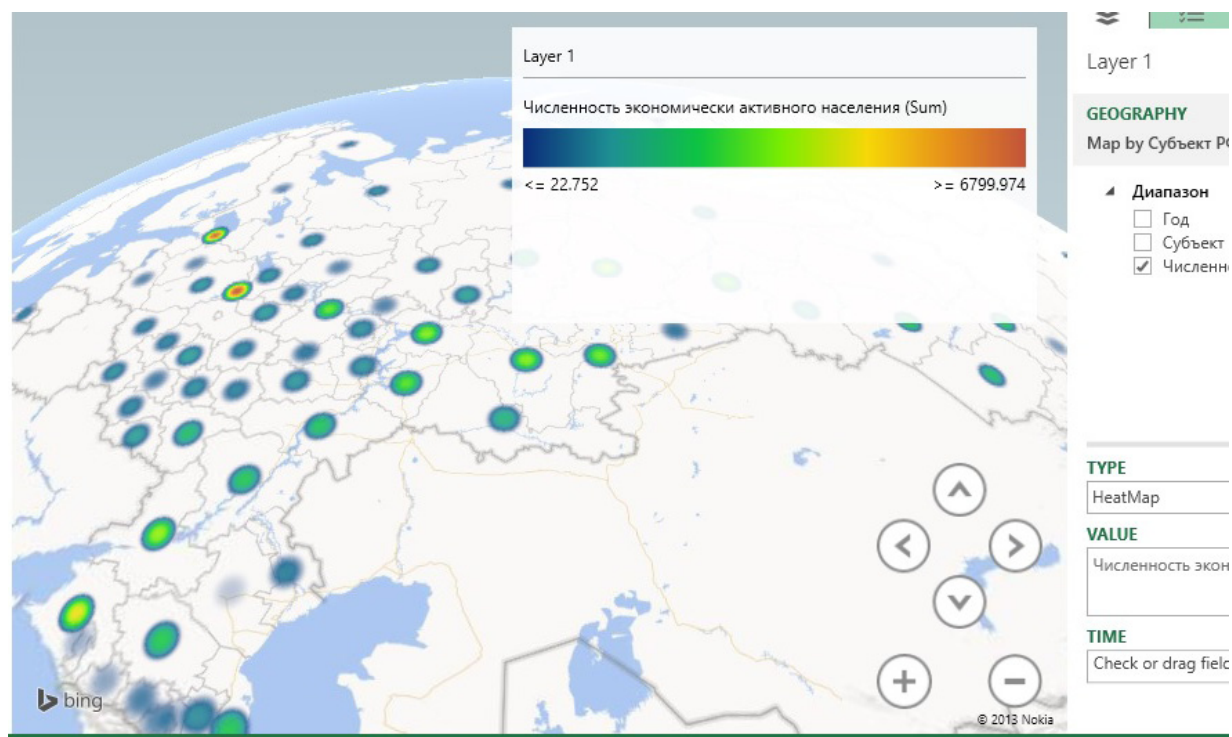


Рис. 5. Столбчатый и пузырьковый вид отображения статистических данных.



Визуализация статистических данных с элементами гис технологий

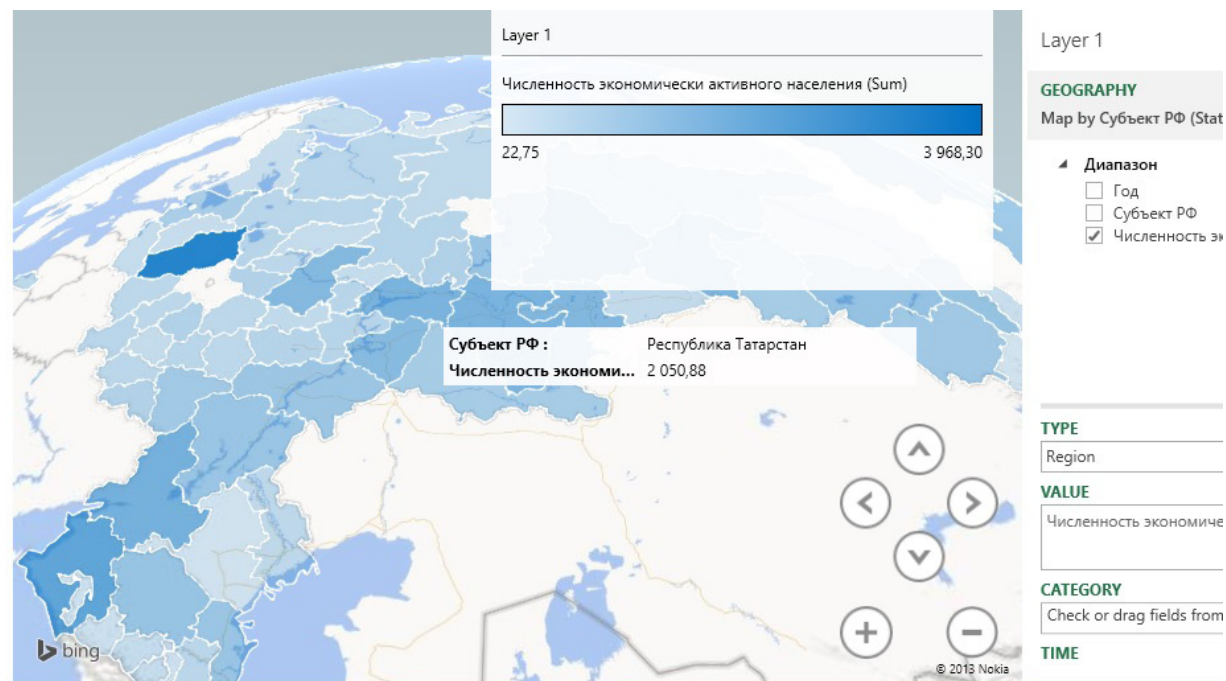


Рис. 6. Тепловая карта и региональный вид отображения статистических данных.

При помощи опции «Themes» можно изменять вид отображения картографической основы. Некоторые темы могут содержать основные дорожные карты, некоторые содержат спутниковые изображения высокой точности. В зависимости от потребностей пользователя можно выбрать наиболее подходящую тему карты.

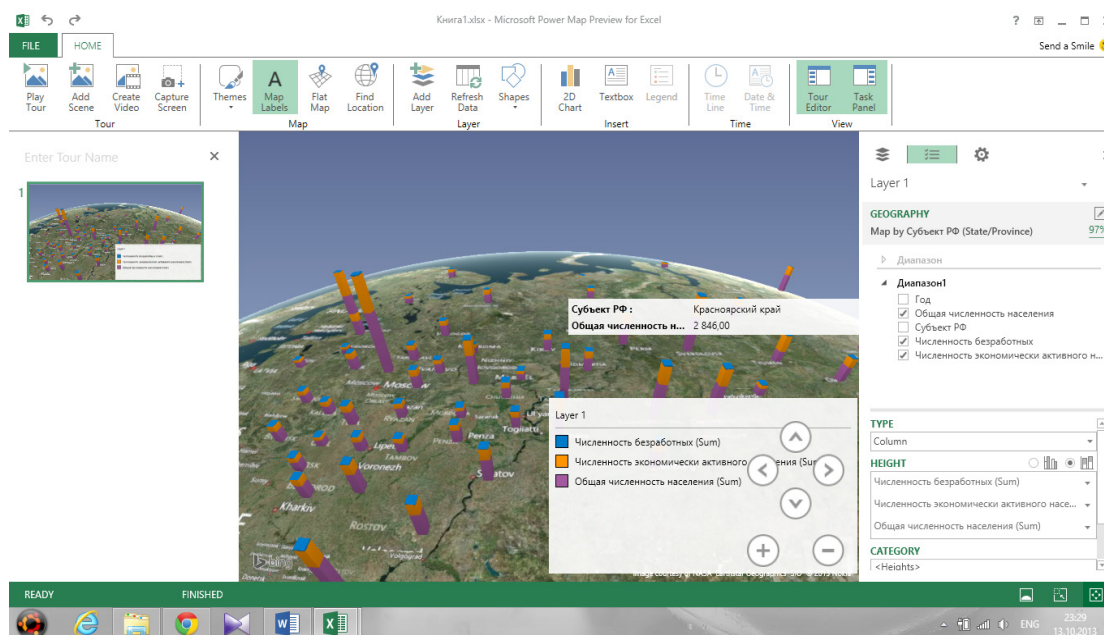


Рис. 7. Спутниковая картоснова.

Инструмент «Flat Map» позволяет спроектировать поверхность Земли со сферы на плоскость, таким образом, чтобы получилась привычная для пользователя карта.

Функция “Find Location” позволяет по адресу, либо наименованию объекта найти его на карте. Причем необходимо отметить, что функция распознает кириллицу, главное, чтобы не было орфографических ошибок.

Для добавления слоя данных аналитики необходимо нажать “Add Layer”.

Если в исходных таблицах происходят какие-либо изменения численных характеристик показателей, следует воспользоваться инструментом “Refresh Data”, который позволяет обновлять отображаемые данные с учетом этих изменений.

Инструменты группы “Insert” позволяют закрепить за туром элементы, уточняющие или конкретизирующие данные аналитики. Кнопка 2D выводит на слой привычную для табличных процессоров 2D-диаграмму или график; при нажатии Textbox можно без привязки к географии вставить какую-либо надпись или комментарий для сопровождения тура; функция “Legend” выводит легенду в соответствии с отображаемыми данными. Нажав правой кнопкой мыши на столбец / пузырек / тепловой элемент карты, и, выбрав “Add Annotation”, можно добавить надпись, привязав её к географическому положению объекта.

Для закрепления слоя необходимо в разделе “Scene Options” присвоить ему наименование, а также подключить визуальный эффект, скорость и длительность его воспроизведения. По полученному отображению данных можно сделать выводы о взаимосвязи численности населения, численности экономически активного населения, числа безработных с географией.

Так, большинство населения России сосредоточено в главной полосе расселения – треугольнике, вершинами которого являются Санкт-Петербург на севере, Сочи на юге и Иркутск на востоке. К северу от этого треугольника благоприятных климатических условий располагается зона тайги и многолетней мерзлоты; к юго-востоку от неё простираются полупустыни и пустыни. В Сибири, площадь которой составляет почти 3/4 территории России, проживает менее четверти населения. Западная и центральная части Европейской России наиболее плотно заселены и урбанизированы. В этих районах расположены крупнейшие города России и традиционные центры культуры и промышленности. На Урале население сконцентрировано главным образом между городами Нижний Тагил и Орск. В Сибири и на Дальнем Востоке население сосредоточено вдоль трассы Транссибирской железной дороги, на которой расположены и её крупнейшие города - Омск, Новосибирск, Красноярск, Иркутск, Хабаровск и Владивосток, а также на

Визуализация статистических данных с элементами гис технологий

территории Кузнецкого угольного бассейна.

Исходя из численности населения по субъектам Российской Федерации распределяется численность экономически активного населения.

Выводы

Инструмент GeoFlow (Microsoft Power Map Preview for Excel) представляет собой эффективное приложение для анализа информации, имеющих пространственную привязку и позволяет наглядно показать изменения, тенденции и закономерности, когда традиционные таблицы и графики не в состоянии этого сделать.

С помощью GeoFlow можно привязывать информацию Excel к местности, исследовать данные в территориальном и хронологическом разрезе, а также выполнять другие операции в интерактивном режиме.

Кроме того, визуализация данных с помощью данной надстройки позволяет конструировать динамические туры (видеоэкскурсии), которые повышают информативность и увеличивают степень понимания и несомненно, дают преимущество перед другими видами визуализации. Добавляя возможности 3D-визуализации данных на карте в виде различных диаграмм с возможностями анимации, табличный процессор Excel получил некоторые отличительные черты ГИС таких как ArcGis или MapInfo.

Таким образом, эта разработка добавляет редактору таблиц еще больше возможностей обработки данных к уже существующим модулям бизнес-аналитики (Business Intelligence) в Excel 2013, таких как Microsoft Data Explorer Preview и Power View.

Не смотря на очевидные преимущества надстройки Microsoft, следует выделить недостатки предварительной версии:

- отсутствие поддержки русского интерфейса;
- отсутствие возможности работы со связанными таблицами (все данные берутся из одной таблицы);
- нет интеграции с SharePoint;
- нет возможности делать срезы пространственных данных (включение / отключение частей диаграмм в режиме просмотра);
- отсутствует возможность свёртки / развёртки (drill-down) категорий слоев или данных;
- ограничения на работу с пространственными данными.

На данный момент аналитическая надстройка GeoFlow имеет весьма неплохие перспективы развития и применения в таких сферах, как бизнес-аналитика, статистический

анализ и прогнозирование, логистика, разработка и принятие управленческих решений во всех структурах (в том числе в органах государственной власти) и т.д.

Библиография:

1. Зиновьев, А. Ю. Визуализация многомерных данных / А.Ю. Зиновьев – Красноярск: Изд. Красноярский ГТУ, 2000. – 180 с.
2. Исакова, О. П. Обработка и визуализация данных физических экспериментов с помощью пакета Origin / Исакова О.П., Тарасевич Ю.Ю., Юзюк Ю.И. – М: Изд. Либроком, 2009. – 138 с.
3. Эйнджел Э. Интерактивная компьютерная графика. Вводный курс на базе OpenGL. –М.: Издательский дом “Вильямс”, 2001. –592 с.
4. Карпик А.П. Методологические и технологические основы геоинформационного обеспечения территорий: Монография. – Новосибирск: СГГА, 2004. – 260 с.
5. Журкин И. Г., Шайтура С. В. Геоинформационные системы. – Москва: КУДИЦ-ПРЕСС, 2009. – 272 с.
6. Питенко А. А. Нейросетевой анализ в геоинформационных системах // Автореферат диссертации на соискание ученой степени кандидата технических наук по специальности 05.13.16, Красноярск, 2000.
7. Инструмент GeoFlow [Официальный сайт]. URL: <http://microsoftbi.ru/tag/geoflow/> (дата обращения: 20.12.2013).
8. Карты Bing [Официальный сайт]. URL: <http://www.bing.com/maps/> (дата обращения: 20.12.2013).
9. Системные требования для GeoFlow [Официальный сайт]. URL: <http://www.microsoft.com/en-us/download/details.aspx?id=38395> (дата обращения: 20.12.2013).
10. Численность экономически активного населения ФСГС [Официальный сайт]. URL: http://www.gks.ru/free_doc/new_site/population/trud/trud1.xls (дата обращения: 20.12.2013)

References:

1. Zinov'ev, A. Yu. Vizualizatsiya mnogomernykh dannykh / A.Yu. Zinov'ev – Krasnoyarsk: Izd. Krasnoyarskii GTU, 2000. – 180 s.
2. Isakova, O. P. Obrabotka i vizualizatsiya dannykh fizicheskikh eksperimentov s pomoshch'yu paketa Origin / Isakova O.P., Tarasevich Yu.Yu., Yuzyuk Yu.I. – M: Izd. Librokom, 2009. – 138 s.
3. Eindhel E. Interaktivnaya komp'yuternaya grafika. Vvodnyi kurs na baze OpenGL. –М.: Izdatel'skii dom “Vil'yams”, 2001. –592 s.
4. Karpik A.P. Metodologicheskie i tekhnologicheskie osnovy geoinformatsionnogo obespecheniya territorii: Monografiya. – Novosibirsk: SGGA, 2004. – 260 s.
5. Zhurkin I. G., Shaitura S. V. Geoinformatsionnye sistemy. – Moskva: KUDITs-PRESS, 2009. – 272 s.

Визуализация статистических данных с элементами ГИС технологий

6. Pitenko A. A. Neurosetevoi analiz v geoinformatsionnykh sistemakh // Avtoreferat dissertatsii na soiskanie uchenoi stepeni kandidata tekhnicheskikh nauk po spetsial'nosti 05.13.16, Krasnoyarsk, 2000.
7. Instrument GeoFlow [Ofitsial'nyi sait]. URL: <http://microsoftbi.ru/tag/geoflow/> (data obrashcheniya: 20.12.2013).
8. Karty Bing [Ofitsial'nyi sait]. URL: <http://www.bing.com/maps/> (data obrashcheniya: 20.12.2013).
9. Sistemnye trebovaniya dlya GeoFlow [Ofitsial'nyi sait]. URL: <http://www.microsoft.com/en-us/download/details.aspx?id=38395> (data obrashcheniya: 20.12.2013).
10. Chislennost' ekonomicheskii aktivnogo naseleniya FSGS [Ofitsial'nyi sait]. URL: http://www.gks.ru/free_doc/new_site/population/trud/trud1.xls (data obrashcheniya: 20.12.2013)