

Visual Mining Studio、Visual R Platform、Deep Learner、Big Data ModuleはAlkanoとしてリニューアルしました。

**横山 暁** 先生／青山学院大学経営学部マーケティング学科 様

## データ分析の研究に、授業に、 R言語を自在に活用

データ分析の領域の研究者として活躍する横山 暁先生。各種学会での研究発表など研究活動だけでなく、大学の教員として後進の指導にも尽力している。その横山先生が研究・教育環境のよりいっそうの充実のために、R言語用統計解析ツール「**Visual R Platform**」(以下**VRP**)を導入した。



青山学院大学  
経営学部 マーケティング学科  
准教授  
横山 暁 先生

### Interview

#### R言語利用の課題をVRPが解決

データ分析の研究で、R言語を活用されていますね。

**横山** 多次元尺度構成法やクラスター分析法によるデータ分析の理論を主な研究テーマとして、マーケティング分野のデータを中心にデータ分析を行い、その結果の検証を行っています。データ分析用のツールやパッケージはいろいろありますが、その中でも近年注目を集めているのが、統計解析向けプログラミング言語「R」(R言語)です。オープンソースのため誰でも無料で使えます。分析に合わせて自分でプログラムを書けるほか、他の研究者が開発したパッケージも利用できるため、私もデータ分析に高い頻度で使用しています。ただ、R言語の運用管理には苦勞していました。データ分析を行う際の前処理の方法、分析プログラムの使用方法、データの入出力の方法など、データを継続的に分析するには、これらのことを分析者が記憶しておく必要があります。そうしないと、分析するデータの生成プロセスが分からなくなったり、分析手法の再利用ができなくなったりします。しかし現実的には全ての面で記憶しておくことは難しく、毎回同じような分析を何度もしてしまうために時間が無駄になる、といった問題がありました。

VRPの印象をお聞かせください。

**横山** 最初に触れたのは学会の展示スペースでした。R言語でやりたい分析が自由にできる支援ツールだという印象を持ちました。どういうデータを、どのように処理して分析するかアイコンでビジュアルに処理でき、ダブルクリックでそれを実行することもできます。また一般的な分析手法から、自分で書いたR言語のプログラムまで利用できます。つまり、分析の効率を高めることも、研究の目的に合わせてコアな解析を行うことも柔軟にできるわけです。さらにR言語を使う際の手間や管理を大幅に減らすことができ、その分、本来の分析作業に時間が割けるようになります。研究の精度や効率を大きく高められそうだと感じ、導入を決意しました。

#### PROFILE

よこやま きたる  
**横山 暁** 先生

慶應義塾大学理工学部管理工学科、同大学院理工研究科開放環境科学専攻 修士課程および後期博士課程修了。博士(工学)。立教大学社会情報教育センター設置準備室(当時)、帝京大学経済学部経営学科を経て、2017年4月青山学院大学経営学部マーケティング学科に着任。専門は多変量解析・データマイニング。特に「単相多元データにおける重複クラスター分析法」に関する研究に従事。「大規模マーケティングデータへのクラスタリング手法の適用」は、平成26-28年度科学研究費助成事業 若手研究(B)に採択。

[裏面に続く]

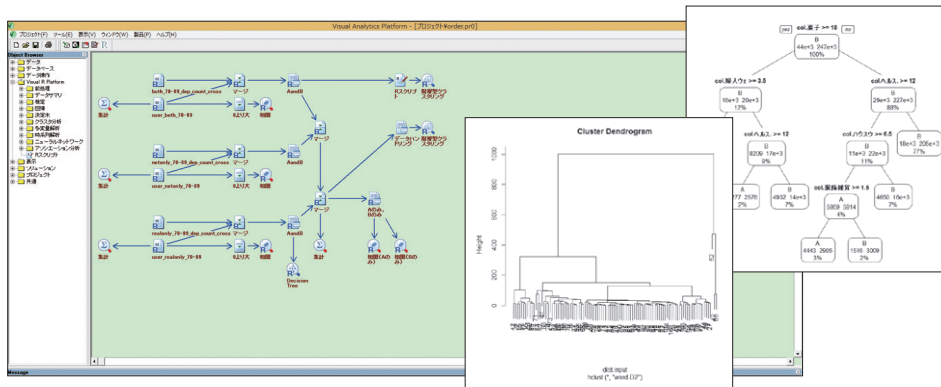


大学の授業やゼミにも活用されているそうですね。

**横山** 自分の研究のために**VRP**を使い始めたのですが、「このツールは学生にも有効だ」とすぐ感じました。R言語はコマンドラインによるCUI環境ですが、それがいわゆる文系の学部にも所属する学生にはハードルを高くする要因になっていました。しかし**VRP**があれば、マウス操作によるGUI環境で分析が進められます。

さらに、このツールで分析した結果は、R言語がインストールされたPC上であれば作業連携ができます。ということは、基本的な分析を大学のPCで**VRP**により行い、授業後は自分のPCにインストールしたR言語やExcelなどでグラフを描くといった作業の続きができます。

授業において特定の分析用ソフトウェアを利用すると、使用できるPCが大学内の一部のPCに限られることが一般的で、自宅で課題を行うことができません。しかし**VRP**とR言語であれば、自宅のPCのR言語環境との作業連携により、授業の構成や難易度をかなり柔軟に設定できるようになります。



**VRP**ならR言語の機能をマウス操作で簡単に使えるため、ツリーの作成やクラスター分析もラクラク行える。

データ解析コンペティションで**VRP**を利用されていますね。

**横山** 経営科学系研究部会連合協議会が毎年開催している「データ解析コンペティション」に参加しています。このコンペティションは産学連携のプロジェクトで、主に企業から提供されたマーケティングデータを用い、学生や社会人が独自の着眼点で分析し、プレゼンテーションを行うものです。私自身も学生時代から参加しており、教員となった今では、ゼミの活動における研究テーマの一つとしてこのコンペティションに参加しています。その分析作業のために、コンペティションに協賛しているNTTデータ数理システムが参加者へ提供している**VRP**のレンタル制度を利用しました。

コンペティションで課題となるデータの規模は年々増加しており、数万件の顧客購買データやウェブアクセスログデータなど、最近では全てのデータをExcelで開くことができないほど膨大になっています。そのため、これまで私のゼミでコンペティションに参加する際には、あらかじめ私がデータの一部を抽出した上で学生に提供し、そこから分析を開始するといった対応をするしかありませんでした。

しかし昨年度は、NTTデータ数理システムが**VRP**をレンタル提供してくださることになったので、データコンペに参加する学生用に導入しました。これにより膨大なデータも容易に開けるようになり、データ全体を見渡してから分析の要点を探ったり、アイデアを出し合ったりすることができるようになりました。本格的な分析作業の“道”を作りやすくなったと感じています。

しかし昨年度は、NTTデータ数理システムが**VRP**をレンタル提供してくださることになったので、データコンペに参加する学生用に導入しました。これにより膨大なデータも容易に開けるようになり、データ全体を見渡してから分析の要点を探ったり、アイデアを出し合ったりすることができるようになりました。本格的な分析作業の“道”を作りやすくなったと感じています。

## 機能をさらに使いこなして研究の幅を広げたい

実際にご利用になっての感想はいかがですか。

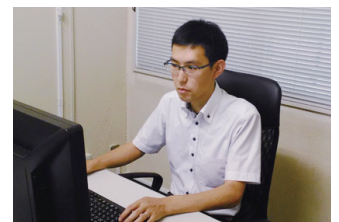
**横山** PCにインストールする段階からサポートがあり、スムーズに使い始めることができました。**VRP**で使用できる分析手法から、自分で書いたプログラムによる独自の分析まで、このツール一つで全て行えるようになり、研究に非常に役立っています。実際に使用していると改善点なども見つけるのですが、NTTデータ数理システムが可能な限り対応してくれます。その際、営業担当の方にポイントを伝えるだけで、開発担当まで正確に伝達されます。スタッフ全員が統計やプログラムに関する知識やスキルを持っているため、信頼感や安心感も高いです。

今後の展望をお聞かせください。

**横山** 私が言うまでもなく、R言語はデータ分析の領域で非常にメジャーな言語になっています。そのR言語を存分に活用

して研究していくための環境を、**VRP**によって充実させることができました。今後はその機能をさらに使いこなして研究の幅を広げていきたいと思っています。

また、学生たちにとって今からR言語に触れておくことは、この先社会に出て仕事をする上で、きっと役に立つはずで**VRP**は、文系の学生にもそれを容易に可能にしてくれます。今後も授業やゼミを通じて学生たちにR言語を体験してもらい、それによって少しでも社会に役立つ人材を送り出していきたいと思っています。



「**VRP**はプロジェクトや途中の集計を保存しておけるのが魅力。R言語だけではできない管理の手間が省け、研究を効率的に進められます」と横山先生。