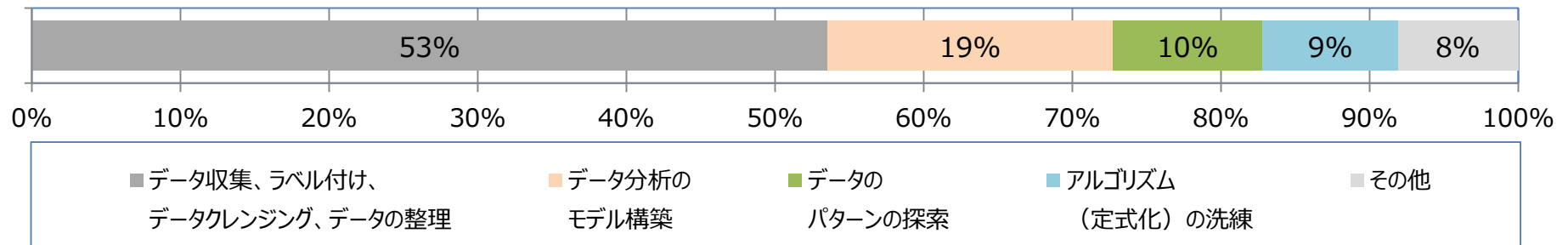


データクレンジングの負担

◆データ分析において、データクレンジング・データ整理は時間がかかり、好まれない作業です。

- 2017年2月における世界のデータサイエンティスト（データ分析者）197名に対する調査では、データクレンジングを含むデータの前処理が最も時間を割いている業務と回答した者が過半の53%となっています。

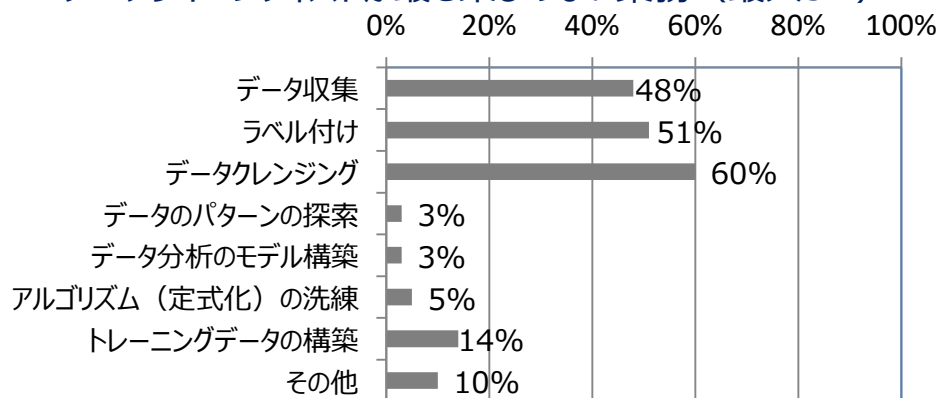
データサイエンティストがデータサイエンスの業務時間で最も時間を割いている業務



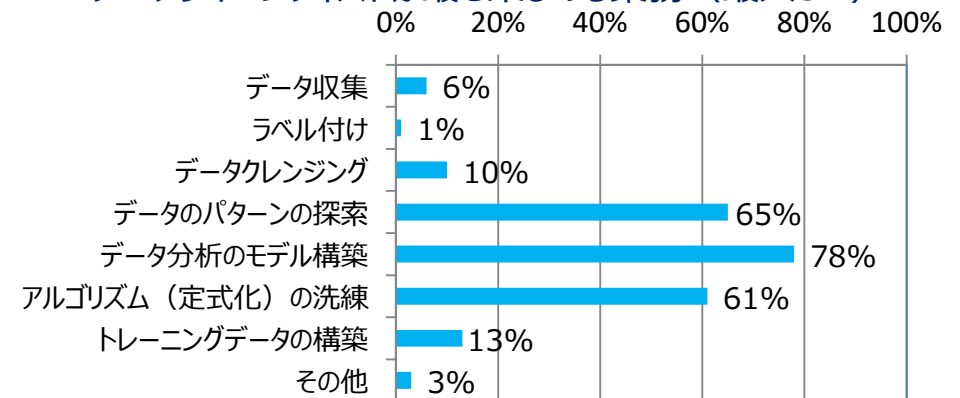
- データクレンジングは、データサイエンティストの業務の中で「最も楽しめない業務」として挙げられています。

- データサイエンティストが最も楽しめる業務として「データ分析のモデル構築」が挙げられています。

データサイエンティストが最も楽しめない業務（最大3つ）



データサイエンティストが最も楽しめる業務（最大3つ）



【出所】2017 Data Scientist Report [CrowdFlower] に基づいて作成

- データクレンジングは労働時間の大部分を占めるとともに心理的な負担になっており、その軽減が課題となっています。