

Top 10

Most important things in Big Data Analysis

1. **Ensure useful quality data**, either by having it or getting it.
2. **Fast access to data** for quick analysis turn-around is key.
3. **All data is flawed**. Make sure you know how and filter it first.
4. **Understand and check input data** and its most basic relations.
5. **Devise a good “measure of goodness”** to track improvements.
6. **Test several methods**, as different methods apply to different cases.
7. **Number of variables**. Typically, +25 does not improve things, but...!
8. **Do not (unknowingly) overtrain methods**. Ensure this by splitting data.
9. **Don't expect miracles from Big Data**. Typical improvements are 5-10%.
10. **Make cross checks**. All data analysis results can be bugged / flawed.