Average Diversity (%) - Average Across All Models (Story)

Statistical test results:
Best VS: VS-Standard
vs Direct: *** (p=0.000)
vs CoT: *** (p=0.000)
vs Sequence: ns (p=0.699)
vs Multi-turn: ** (p=0.005)

Bars:
- Direct: 11.1±1.3
- CoT: 11.6±1.6
- Sequence: 19.0±2.4
- Multi-turn: 13.9±2.4
- VS-Standard: 18.5±3.1
- VS-CoT: 20.1±1.8
- VS-Multi: 17.8±3.0